



**Next-Generation Sequencing Identifies  
Mechanisms Of Tumourigenesis Caused By Loss  
Of *SMARCB1* In Malignant Rhabdoid Tumours**

**Martina Anna Finetti**

Thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

**Newcastle University**

**Faculty of Medical Sciences**

**Northern Institute for Cancer Research**

December 2014

## **Declaration**

I certify that no part of the material documented in this thesis has previously been submitted for a degree or other qualification in this or any other university. I declare that this thesis represents my own unaided work, carried out by myself, except where it is acknowledged otherwise in the thesis text.

Martina Finetti

December 2014

## **Acknowledgements**

I would like to express my sincere gratitude to my supervisors, Dr. Daniel Williamson, Professor Steven Clifford, and Professor Simon Bailey. Their continuous support and their extraordinary knowledge of paediatric cancer and bioinformatics have helped me immensely throughout my research.

I would also like to thank my assessors Professor John Lunec and Dr. Luke Gaughan and Professor Craig Robson for their constant guidance and encouragement for the last four years.

I would like to thank the members of the Paediatric Brain Tumour Research Group. In particular I would like to thank Dr. Rebecca Hill, Matthew Selby and Dr Dolores Hamilton for the support, the encouragement and the time spent together in good and bad times.

Special thanks go to all people at Sir James Spence Institute for their technical help and support. I would in particular thank Stefano and Alicia for our friendship: thank you for the moral support and the all the laughs we shared in between experiments!

I would like to thank the NECCR, Children with cancer UK, Brain Trust, Love Oliver, CCLG, Karen and Iain Wark, The Smiley Ridley Fund, whose financial support made this project possible.

Finally, I would like to thank my family: my Mum Chiara, Lorenzo, Giulia and Al. Thank you for supporting me especially in the bad times and for all your caring: without your encouragement I would not have made it.

My final thanks are reserved for my Dad Roberto: your love for knowledge inspires me to continue my studies and accomplish my dreams.

## **Dedication**

This thesis is dedicated to the memory of my Dad, Roberto Finetti.

*How I wish, how I wish you were here.*

*We're just two lost souls swimming in a fish bowl, year after year,*

*Running over the same old ground.*

*What have we found?*

*The same old fears.*

*Wish you were here*

*-Pink Floyd-*



## Abstract

**Introduction:** Malignant Rhabdoid Tumours (MRT) are unique malignancies caused by biallelic inactivation of a single gene (*SMARCB1*). *SMARCB1* encodes for a protein that is part of the SWI/SNF chromatin remodelling complex, responsible for the regulation of hundreds of downstream genes/pathways. Despite the simple biology of these tumours, no studies have identified the critical pathways involved in tumorigenesis. The understanding of downstream effects is essential to identifying therapeutic targets that can improve the outcome of MRT patients.

**Methods:** RNA-seq and 450K-methylation analyses have been performed in MRT human primary malignancies (n > 39) and in 4 MRT cell lines in which lentivirus was used to re-express *SMARCB1* (G401, A204, CHLA-266, and STA-WT1). The MRT cell lines were treated with 5-aza-2'-deoxycytidine followed by global gene transcription analysis (RNA-seq and 450K-methylation) to investigate how changes in methylation lead to tumorigenesis.

**Results:** We show that primary Malignant Rhabdoid Tumours present a unique and distinct expression/methylation profile which confirms that MRT broadly constitute a single and different tumour type from other paediatric malignancies. However, despite their common cause MRT can be sub-grouped by location (i.e. CNS or kidney). We observe that re-expression of *SMARCB1* in MRT cell lines determines activation/inactivation of specific downstream pathways such as IL-6/TGF beta. We also observe a direct correlation between alterations in methylation and gene expression in *CD44*, *GLI2*, *GLI3*, *CDKN1A*, *CDKN2A* and *JARID* after *SMARCB1* re-expression. Loss of *SMARCB1* also promotes expression of aberrant isoforms and novel transcripts and causes genome-wide changes in SWI/SNF binding.

**Conclusion:** Next generation transcriptome and methylome analysis in primary MRT and in functional models give us detailed downstream effects of *SMARCB1* loss in Malignant Rhabdoid Tumours. The integration of data from both primary and functional models has provided, for the first time, a genome-wide catalogue of *SMARCB1* tumourigenic changes (validated using systems biology). Here we show how a single

deletion of *SMARCB1* is responsible for deregulation of expression, methylation status and binding at the promoter regions of potent tumour-suppressor genes. The genes, pathways and biological mechanisms indicated as key in tumour development may ultimately be targetable therapeutically and will lead to better treatments for what is currently one of the most lethal paediatric cancers.

## List of abbreviation

<b>H2A</b>	Histone 2A
<b>μg</b>	Microgram
<b>μl</b>	Microliter
<b>3' UTR</b>	Three Prime Untranslated Region
<b>5-azaCdR</b>	5-Aza-2'-Deoxycytidine
<b>A</b>	Adenine
<b>AATK</b>	Apoptosis-Associated Tyrosine Kinase
<b>Ab</b>	Antibody
<b>ABR</b>	Active BCR-Related
<b>AKT</b>	V-Akt Murine Thymoma Viral Oncogene
<b>AML</b>	Acute Myelogenous Leukaemia
<b>APC</b>	Adenomatous Polyposis Coli
<b>ARID1A/ BAF250a</b>	AT Rich Interactive Domain 1A (SWI-Like)

<b>ATPase</b>	Adenylpyrophosphatase
<b>ATRT</b>	Atypical Teratoid Rhabdoid Tumours
<b>AZA</b>	5-Aza-2-Deoxycytidine
<b>B2M</b>	Beta-2-Microglobulin
<b>BAF155</b>	BRG1-Associated Factor 155
<b>BAF170</b>	BRG1-Associated Factor 170
<b>BASH</b>	Beadarray Subversion Of Harshlight
<b>BCA</b>	Bicinchoninic Acid Assay
<b>BCL2</b>	B-Cell CLL / Lymphoma 2
<b>B-FABP</b>	Brain Fatty Acid-Binding Protein
<b>BHLHE40</b>	Basic Helix-Loop-Helix Family, Member E40
<b>BID</b>	BH3 Domain-Containing Proapoptotic Bcl2 Family Member
<b>BK</b>	PDZ-Binding Kinase
<b>BLAST</b>	Basic Local Alignment Search Tool

<b>BM1</b>	Polycomb Complex Protein
<b>BMI1</b>	BMI Polycomb Ring Finger Oncogene
<b>BP</b>	Base Pair
<b>BRAF</b>	V-Raf Murine Sarcoma Viral Oncogene Homologue B1
<b>BRCA1/2</b>	Breast Cancer 1/2
<b>BRG1</b>	Brahma-Related Gene 1
<b>BRM</b>	Brahma
<b>BTG4</b>	B-Cell Translocation Gene 4
<b>C</b>	Cytosines
<b>C2</b>	Curated Gene Sets
<b>C6</b>	Oncogenic Signatures
<b>CASP8</b>	Caspase 8, Apoptosis-Related Cysteine Peptidase
<b>CCDC46</b>	Coiled-Coil Domain Containing 46
<b>CCDC74B</b>	Coiled-Coil Domain Containing 74B
<b>CCLG</b>	Children's Cancer And Leukaemia Group

<b>CCNB1</b>	Cyclin B1
<b>CD8</b>	(Cluster Of Differentiation 8
<b>CDH1</b>	Cadherin 1, Type 1, E-Cadherin (Epithelial)
<b>CDKN1A</b>	Cyclin Dependent Kinase Inhibitor 1A
<b>CDKN2A</b>	Cyclin-Dependent Kinase Inhibitor 2A
<b>cDNA</b>	Complementary DNA
<b>CDS</b>	Similarly Coding Sequences
<b>CHD5</b>	Chromo domain Helicase DNA Binding Protein 5
<b>ChIP</b>	Chromatin Immunoprecipitation
<b>ChIP-seq</b>	Chromatin Immunoprecipitation Sequencing
<b>CLAS</b>	Classic Type Medulloblastoma
<b>CML</b>	Chronic Myelogenous Leukaemia
<b>CNAs</b>	Copy-Number Aberrations
<b>CNN</b>	Copy Number Neutral
<b>CNS</b>	Central Nervous System

<b>COL1A2</b>	Collagen, Type I, Alpha 2
<b>Conc</b>	Concentration
<b>CpG</b>	Cytosine-Guanine Dinucleotide
<b>cRNA</b>	Complementary RNA
<b>CTCF</b>	CCCTC-Binding Factor (Zinc Finger Protein)
<b>CTL</b>	Control
<b>CTTNB1</b>	Catenin (Cadherin-Associated Protein), Beta 1
<b>Cy3</b>	Cyanine 3 Dye
<b>Cy5</b>	Cyanine 5 Dye
<b>DAVID</b>	Discovery Database For Annotation, Visualization And Integrated
<b>DAZ</b>	Deleted In Azoospermia
<b>DCLK1</b>	Doublecortin-Like Kinase 1
<b>DHH</b>	Desert Hedgehog
<b>DLC1</b>	Deleted In Liver Cancer 1
<b>DMSO</b>	Dimethyl Sulfoxide

<b>DN</b>	Dominant Negative
<b>DN</b>	Desmoplastic / Nodular Medulloblastoma
<b>DNA</b>	Deoxyribose Nucleic Acid
<b>DNA2</b>	DNA Replication Helicase/Nuclease 2
<b>DNM3OS</b>	NM3 Opposite Strand/Antisense RNA
<b>DNMT1</b>	DNA Methyltransferase 1
<b>DNMT3A</b>	DNA Methyltransferase 3A
<b>DNMT3B</b>	DNA Methyltransferase 3B
<b>dNTPs</b>	Nucleoside Triphosphates Containing Deoxyribose
<b>DRAXIN</b>	Dorsal Inhibitory Axon Guidance Protein
<b>E. coli</b>	Escherichia Coli
<b>E2F1</b>	E2F Transcription Factor 1
<b>E2F3</b>	E2F Transcription Factor 3
<b>E2F4</b>	E2F Transcription Factor 4
<b>ECRT</b>	Extra Cranial Rhabdoid Tumours



<b>EGFR</b>	Epidermal Growth Factor Receptor
<b>EGL</b>	External Granule Layer
<b>EPCAM</b>	Epithelial Cell Adhesion Molecule
<b>ER-</b>	Estrogen-Receptor Negative
<b>ERBB2</b>	V-Erb-B2 Erythroblastic Leukaemia Viral Oncogene Homologue2,Neuroma / Glioblastoma
<b>ES</b>	Embryonic Stem
<b>ES</b>	Enrichment Score
<b>EU-RHAB</b>	European Rhabdoid Registry
<b>EWAS</b>	Epigenome-Wide Association Studies
<b>EWS</b>	Ewing Sarcoma Breakpoint Region 1
<b>EZH2</b>	Enhancer Of Zeste Homolog 2
<b>FBXW8</b>	F-Box And WD Repeat Domain Containing 8,
<b>FDR</b>	False Discovery Rate
<b>FPKM</b>	Fragments Per Kilo-Base Of Exon Per Million Fragments Mapped)

<b>FPKM</b>	Fragments Per Kilobases Per Millon Reads
<b>Fwd</b>	Forward
<b>G</b>	Guanine
<b>GA</b>	Genome Analyser
<b>Gal-1</b>	Galactokinasi 1
<b>GALNT2</b>	Polypeptide N-Acetylgalactosaminyltransferase 2
<b>GAPDH</b>	Glyceraldehyde-3-Phosphate Dehydrogenase
<b>GC%</b>	Guanine-Cytosine Content
<b>GCS</b>	Genome Analyser Control Software
<b>gDNA</b>	Genomic DNA
<b>GEO</b>	Gene Expression Omnibus
<b>GFAP</b>	Glial Fibrillary Acidic Protein
<b>GLI1</b>	GLI Family Zinc Finger 1
<b>GLI2</b>	GLI Family Zinc Finger 2
<b>GLI2</b>	GLI Family Zinc Finger 2

<b>GLI3</b>	GLI Family Zinc Finger 3
<b>GLI3</b>	GLI Family Zinc Finger 3
<b>GPx-1</b>	Glutathione Peroxidase 1
<b>GPx-2</b>	Glutathione Peroxidase 2
<b>GSEA</b>	Gene Set Enrichment Analysis
<b>GTP</b>	Guanine Triphosphate
<b>Gy</b>	Gray (Unit)
<b>g</b>	G-Force
<b>H2.AZ</b>	Variant Histone 2 Az
<b>H2B</b>	Histone 2B
<b>H2O</b>	Dihydrogen Monoxide
<b>H3</b>	Histone 3
<b>H3K27me3</b>	Histone 3 Lysine 27 Trimethylation
<b>H3K36me3</b>	Histone 3 Lysine 36 Trimethylation
<b>H3K4me3</b>	Histone 3 Lysine 4 Trimethylation

<b>H4</b>	Histone 4
<b>HATs</b>	Histone Acetyltransferases
<b>HC</b>	Hierarchical Clustering
<b>HC</b>	Hierarchical Clustering
<b>HDAC1</b>	Histone Deacetylase 1
<b>HDAC2</b>	Histone Deacetylase 2
<b>HDACs</b>	Histone Deacetylase
<b>HDCT</b>	High-Dose-Chemotherapy
<b>HDMs</b>	Histone Demethylases
<b>HepG2</b>	Human Liver Hepatocellular Carcinoma Cell Line
<b>HER2-</b>	Human Epidermal Growth Factor Negative
<b>HFE</b>	Hemochromatosis
<b>HG19</b>	Human Genome 19
<b>HGF</b>	Hepatocyte Growth Factor
<b>HGP</b>	Human Genome Project

<b>Hh</b>	Hedgehog
<b>HIC1</b>	Hypermethylated In Cancer 1
<b>HMGA2</b>	High Mobility Group AT-Hook 2
<b>HMTs</b>	Histone Methyltransferases
<b>HOPX</b>	HOP Homeobox
<b>HOXA9</b>	Homeobox A9
<b>HOXD3</b>	Homeobox D3
<b>ICGC</b>	Cancer Genome Consortium
<b>IDH1</b>	Isocitrate Dehydrogenase 1
<b>IgG</b>	Immunoglobulin G
<b>IGV</b>	Integrative Genomics Viewer
<b>IL1</b>	Interleukin 1, Beta
<b>IL17RE</b>	Interleukin 17 Receptor E
<b>IL6</b>	Interleukin 6
<b>IMDM</b>	Iscove's Modified Dulbecco's Media

<b>Infl</b>	Infinium I
<b>InflII</b>	Infinium II
<b>INO80</b>	Inositol Requiring 80
<b>IP</b>	Immunoprecipitated
<b>IPA®</b>	Ingenuity Pathway Analysis
<b>IQR</b>	Interquartile Range
<b>IRF1</b>	Interferon Regulatory Factor 1
<b>ISWI</b>	Imitation Switching Defective
<b>ITIH2</b>	Inter-Alpha (Globulin) Inhibitor H2
<b>JARID2</b>	Jumonji, AT Rich Interactive Domain 2,
<b>KRAS</b>	V-Ki-Ras2 Kirsten Rat Sarcoma Viral Oncogene Homolog A
<b>LB</b>	Luria-Bertani Media
<b>LBX1</b>	Ladybird Homeobox 1
<b>LCA</b>	Large Cell / Anaplastic Medulloblastoma

<b>LM-PCR</b>	Ligation Mediated Polymerase Chain Reaction
<b>MAX</b>	MYC Associated Factor X
<b>MB</b>	Medulloblastoma
<b>MBEN</b>	Medulloblastoma With Extensive Nodularity
<b>MBP</b>	Methyl-Cytosine Binding Protein
<b>MECP2</b>	Methyl Cpg Binding Protein 2
<b>MgCl<sub>2</sub></b>	Magnesium Chloride
<b>MGMT</b>	O(6)-Methylguanine-DNA Methyltransferase
<b>MGMT</b>	O-6-Methylguanine-DNA Methyltransferase
<b>MLL</b>	Lysine (K)-Specific Methyltransferase 2A
<b>mM</b>	MilliMolar
<b>MOI</b>	Multiplicity Of Infection
<b>mRNA</b>	Messenger RNA
<b>MSI1</b>	Musashi RNA-Binding Protein 1

<b>MTS</b>	(3- (4, 5-Dimethylthiazol-2-Yl) -5- (3-Carboxymethoxyphenyl) -2 - (4-Sulfophenyl) -2H-Tetrazolium)
<b>mV</b>	Millivolts
<b>MYCC</b>	V-Myc Myelocytomatosis Viral Oncogene Homologue (Avian)Viii
<b>MYCN</b>	V-Myc Myelocytomatosis Viral Related Oncogene, Neuroblastoma Derived (Avian)
<b>MYLK</b>	Myosin Light Chain Kinase
<b>NaCl</b>	Sodium Chloride
<b>NADH</b>	Nicotinamide Adenine Dinucleotide
<b>NADPH</b>	Nicotinamide Adenine Dinucleotide Phosphate
<b>NaH<sub>2</sub>PO<sub>4</sub></b>	Sodium Phosphate Monobasic Monohydrate,
<b>NDC80/HEC1</b>	Kinetochore Protein Hec1
<b>NDRs</b>	Nucleosome-Depleted Regions
<b>NES</b>	Normalised Enrichment Scores
<b>NEUROG1</b>	Neurogenin 1



<b>ng</b>	Nanogram
<b>NGS</b>	Next Generation Sequencing
<b>NMB</b>	Newcastle Medulloblastoma
<b>NMF</b>	Non-Negative Matrix Factorisation
<b>NOTCH4</b>	Neurogenic Locus Notch Homolog 4
<b>NPTX2</b>	Neuronal Pentraxin II
<b>NTN4</b>	Netrin 4
<b>NuRD/Mi-2/CHD</b>	Nucleosome Remodelling And Deacetylation /Chromo-Domain, Helicase, DNA Binding
<b>OS</b>	Overall Survival
<b>OTX2</b>	Orthodenticle Homeobox 2
<b>P13K</b>	Phosphoinositide 3-Kinase
<b>padj</b>	P Value Adjusted
<b>PATCH1</b>	Patched 1
<b>PCA</b>	Principal Component Analysis

<b>PCA</b>	Principle Component Analysis
<b>PCAM</b>	Cell Adhesion Molecule
<b>PCDH</b>	Pcdh-CMV-MCS-EF1-Puro
<b>PcG</b>	Polycomb Group Proteins
<b>PcGs</b>	Polycomb Group Proteins
<b>PCR</b>	Polymerase Chain Reaction
<b>PDE4D</b>	Phosphodiesterase 4D, Camp-Specific
<b>PDGF</b>	Platelet Derived Growth Factor
<b>PGAM1</b>	Phosphoglycerate Mutase 1
<b>PHF11</b>	PHD Finger Protein 11
<b>PK</b>	Protein Kinases
<b>pM</b>	Pico Molar
<b>PNET</b>	Primitive Neuroectodermal Tumor
<b>Pol III</b>	Rna Polymerase Iii
<b>Pol II</b>	Rna Polymerase Ii

<b>Poly(I): Poly(C)</b>	Polyinosinic–polycytidylic acid
<b>PR-</b>	Progesteron Receptor Negative
<b>PRC1</b>	Polycomb Repressive Complex 1
<b>PRC2</b>	Polycomb Repressive Complex 2
<b>PRDM16</b>	PR Domain Containing 16
<b>prop</b>	Proportion
<b>PTCH1</b>	Patched 1
<b>PTEN</b>	Phosphatase And Tensin Homolog
<b>OCT4</b>	Octamer-Binding Transcription Factor 4
<b>QC</b>	Quality Control
<b>q-PCR</b>	Quantitative Polymerase Chain Reaction
<b>RAP1</b>	Ras-Proximate-1
<b>RB1</b>	Retinoblastoma 1
<b>Rev</b>	Reverse
<b>RHOGAP</b>	Rho GTPase Activating protein

<b>RIN</b>	RNA Integrity Number
<b>RNA</b>	Ribonucleic Acid
<b>RNF220</b>	Ring Finger Protein 220
<b>rpm</b>	Revolutions Per Minute
<b>RPMI</b>	Roswell Park Memorial Institute Medium
<b>RPS27L</b>	Ribosomal Protein S27-Like
<b>rRNA</b>	Ribosomal RNA
<b>RT</b>	Reverse Transcriptase
<b>RTA</b>	Real Time Analysis
<b>RT-PCR</b>	Reverse Transcriptase PCR
<b>SALL4</b>	Spalt-Like Transcription Factor 4
<b>SBS</b>	Sequencing-By-Synthesis
<b>SCS</b>	Sequencing Control Software
<b>SDS</b>	Sodium Dodecyl Sulphate
<b>SERPINF1</b>	Pigment Epithelium-Derived Factor

<b>SHH</b>	Sonic Hedgehog
<b>SOX1</b>	Sex Determining Region Y-Box 1)
<b>SOX2</b>	SRY (Sex Determining Region Y)-Box 2
<b>ST18</b>	Suppression Of Tumorigenicity 18
<b>ST5</b>	Suppression Of Tumorigenicity 5
<b>STAR</b>	Transcripts Alignment And Reconstruction Tool
<b>STAT3</b>	Signal Transducer And Activator Of Transcription 3
<b>Subunit-MTA1</b>	Metastasis –Associated Gene 1
<b>SUFU</b>	Suppressor Of Fused Homolog
<b>SUSD2</b>	Sushi Domain Containing 2
<b>SWAN</b>	Subset-Quantile Within Array Normalisation
<b>SWI/SNF</b>	Switching Defective /Sucrose Non-Fermenting
<b>T</b>	Thymines
<b>TBFB1</b>	Transforming Growth Factor, Beta 1
<b>TBP</b>	TATA-Binding Protein

<b>TE</b>	Tris EDTA Buffer
<b>TEs</b>	Transposable Elements
<b>T<sub>m</sub></b>	Melting Temperature
<b>TNXB</b>	Tenascin XB
<b>TOPK</b>	T-LAK-Cell-Originated Protein Kinase
<b>TP53</b>	Tumor Protein P53
<b>TRIS</b>	Tris (Hydroxymethyl) Aminomethane
<b>TrxGs</b>	Trithorax Group Proteins
<b>TSSs</b>	Transcription Start Sites
<b>TTBS</b>	Tween 20 Tris-Buffered Saline
<b>TUBA1C</b>	Tubulin Alpha-1C Chain
<b>U</b>	Uracils
<b>Un</b>	Untreated
<b>Vol</b>	Volume
<b>vst</b>	Variance Stabilising Transform

<b>WHO</b>	World Health Organization
<b>WT</b>	Wild Type
<b>WT1</b>	Wilms Tumor 1
<b>YAP1</b>	Yes-Associated Protein 1
<b>ZSCAN4</b>	Zinc Finger And SCAN Domain Containing 4

## List of content

<b>Declaration.....</b>	<b>I</b>
<b>Acknowledgements.....</b>	<b>II</b>
<b>Dedication .....</b>	<b>III</b>
<b>Abstract.....</b>	<b>IV</b>
<b>List of abbreviation .....</b>	<b>VI</b>
<b>List of content .....</b>	<b>XXVII</b>
<b>List of figures .....</b>	<b>XLII</b>
<b>List of tables.....</b>	<b>LVIII</b>
<b>Chapter 1 Introduction .....</b>	<b>1</b>
<b>1.1 Atypical Teratoid Rhabdoid Tumours (ATRT).....</b>	<b>2</b>
1.1.1 Epidemiology .....	4
1.1.2 Genetic epidemiology .....	6
1.1.3 Treatment .....	9
1.1.4 Extra Cranial Rhabdoid Tumours (ECRT).....	11
1.1.5 Rhabdoid predisposition syndrome in CNS tumours and tumours with SMARCB1 germ line mutation .....	12
1.1.5.1 Biological and molecular differences between Atypical Teratoid Malignant Rhabdoid Tumours and other embryonal tumours of the CNS.....	13



<b>1.2</b>	<b>Epigenetic changes in cancer .....</b>	<b>17</b>
1.2.1	DNA methylation .....	18
1.2.2	Changes in histone modification.....	20
1.2.3	Chromatin remodelling complexes .....	25
<b>1.3</b>	<b>SWI/SNF complexes .....</b>	<b>28</b>
1.3.1	Role of SWI/SNF complexes in transcriptional regulation.....	30
1.3.1.1	<i>The critical role of the SWI/SNF complex in cancer development.....</i>	<i>32</i>
1.3.2	SWI/SNF subunits and tumourigenesis .....	34
1.3.2.1	<i>SWI/SNF subunits are frequently mutated in a variety of tumours .....</i>	<i>34</i>
1.3.2.2	<i>The ATPase subunits of SWI/SNF complex: role in tumourigenesis.....</i>	<i>36</i>
1.3.2.3	<i>The role of SMARCB1 in Malignant Rhabdoid Tumours.....</i>	<i>37</i>
1.3.2.4	<i>SMARCA4 subunit in SMARCB1 deficient cells: loss of a delicate balance.....</i>	<i>40</i>
<b>1.4</b>	<b>Next generation sequencing and array-based analysis techniques: new techniques in tumour characterisation .....</b>	<b>41</b>
1.4.1	Next generation sequencing .....	41
1.4.2	Transcriptome profiling: RNA sequencing .....	42
1.4.3	DNA-protein interaction: ChIP sequencing .....	45
1.4.4	Array-based analysis techniques: Methylation array .....	47
<b>1.5</b>	<b>Summary and Aims .....</b>	<b>49</b>

<b>Chapter 2 Materials and Methods .....</b>	<b>52</b>
<b>2.1 Primary tumour cohorts .....</b>	<b>53</b>
<b>2.2 Cell lines .....</b>	<b>55</b>
2.2.1 Cell line cultures .....	55
2.2.2 Proliferation assay .....	55
2.2.3 Treatment of Malignant Rhabdoid tumours cell lines with 5-aza-2'-deoxycytidine .....	57
<b>2.3 SDS-PAGE and Western Blot.....</b>	<b>57</b>
2.3.1 Protein extraction and quantification .....	57
2.3.2 Protein electrophoresis and blotting.....	57
<b>2.1 Expression quantitation by real time PCR.....</b>	<b>59</b>
2.1.1 DNA/RNA extraction, purification .....	59
2.1.1.1 Phenol/chloroform extraction of DNA .....	60
2.1.2 DNA/RNA quantification and quality assessment .....	60
2.1.3 Oligonucleotides .....	61
2.1.4 Quantitative polymerase chain reaction .....	62
2.1.4.1 Relative quantification of gene expression data using the $\Delta\Delta CT$ method .....	62
<b>2.2 Lentiviral Production and Infection .....</b>	<b>63</b>
2.2.1 Determining viral titre.....	63

2.2.2	Multiply of infection .....	64
2.2.3	Stable lentiviral infection of Rhabdoid cells .....	64
<b>2.3</b>	<b>Bacterial strains .....</b>	<b>64</b>
2.3.1	Plasmids .....	65
2.3.1.1	<i>PCDH-CMV-MCS-EF1-Puro</i> .....	65
2.3.1.2	<i>SMARCB1 wild type</i> .....	65
2.3.1.3	<i>pBABE SMARCA4 wild type and dominant negative</i> .....	65
2.3.1.4	<i>Packaging plasmids</i> .....	66
2.3.2	Bacterial culture .....	66
2.3.3	Harvesting of bacteria from broth cultures and extraction of nucleic acids .....	67
2.3.4	Quantification of nucleic acids by spectroscopy.....	67
2.3.5	DNA manipulation techniques .....	67
2.3.5.1	<i>DNA restriction</i> .....	67
2.3.5.2	<i>Agarose electrophoresis</i> .....	68
2.3.5.3	<i>Purification of digested DNA</i> .....	68
2.3.5.4	<i>Oligonucleotides</i> .....	69
2.3.5.5	<i>Polymerase chain reaction (PCR)</i> .....	70
2.3.6	Cloning.....	71

2.3.6.1	<i>Cloning PCR amplicons</i> .....	71
2.3.6.2	<i>Transformation</i> .....	71
2.3.6.3	<i>Blunt and cohesive ended ligation</i> .....	71
2.3.6.4	<i>TA Cloning</i> .....	72
2.3.6.5	<i>Site-directed mutagenesis: C2979T SMARCA4</i> .....	73
2.3.6.6	<i>Sequencing</i> .....	73
<b>2.4</b>	<b>Chromatin Immunoprecipitation</b> .....	<b>75</b>
2.4.1	Chromatin enrichment determination .....	76
2.4.1.1	<i>Data analysis and normalisation of ChIP-qPCR data: Percent input method</i> .....	76
<b>2.5</b>	<b>Assessment of gene methylation status</b> .....	<b>77</b>
2.5.1	Assessment of 5-aza-2'deoxyctidine-induced demethylation .....	77
2.5.2	Bisulphite conversion and sequencing .....	78
<b>Chapter 3 Profiling and Bioinformatics Strategy</b> .....		<b>82</b>
<b>3.1</b>	<b>Introduction</b> .....	<b>83</b>
<b>3.2</b>	<b>RNA sequencing pipeline</b> .....	<b>83</b>
3.2.1	Samples preparation .....	84
3.2.2	Library preparation .....	84
3.2.3	Solexa sequencing (Illumina Hi-seq 2000) .....	84
3.2.4	Primary data processing .....	87

3.2.4.1	<i>Raw data quality control</i> .....	87
	<i>Sequence Alignment (Star Aligner) RNA-STAR: ultrafast universal RNA-seq aligner</i> .....	89
3.2.4.2	<i>Alignment Quality Control (SeqQC)</i> .....	89
3.2.5	Differential Expression (DEseq) .....	90
3.2.6	Differential Exon Usage (DEX-seq) .....	91
3.2.7	Transcriptome assembly: differential CDS, promoter usage, TSS usage and novel isoforms .....	92
3.2.8	Transcript fusion detection.....	95
3.2.9	Identification of viral integration on human genome.....	97
<b>3.3</b>	<b>450K Chip Analysis Methylation Pipeline.....</b>	<b>98</b>
3.3.1	HumanMethylation450 BeadChip assays .....	98
3.3.2	Primary data processing.....	100
3.3.2.1	<i>Methylation Microarray Quality Control</i> .....	101
3.3.2.2	<i>Subset-quantile Within Array Normalisation (SWAN) to correct InfI/InfII shift</i> .....	102
3.3.3	Differential methylation pattern.....	104
3.3.4	Detection of Copy-Number Aberrations (CNAs) .....	104
<b>3.4</b>	<b>Unsupervised cluster analysis.....</b>	<b>105</b>
3.4.1	Hierarchical clustering .....	105

3.4.2	Partitioning methods: k-means.....	108
3.4.3	Principal component analysis.....	108
3.4.4	Non-negative matrix factorisation .....	110
3.4.5	Consensus-based unsupervised clustering approach.....	112
<b>3.5</b>	<b>Gene pathway analysis .....</b>	<b>112</b>
3.5.1	Gene Set Enrichment Analysis (GSEA).....	112
3.5.2	Ingenuity pathway analysis (IPA®) .....	113
3.5.3	Discovery Database for Annotation, Visualisation and Integrated (DAVID) .....	114
 <b>Chapter 4 Establishing the Role of <i>SMARCB1</i> Loss in Malignant Rhabdoid Tumour Cells.....</b>		
<b>4.1</b>	<b>Introduction .....</b>	<b>116</b>
<b>4.2</b>	<b>Aims .....</b>	<b>119</b>
<b>4.3</b>	<b>Re-expression of <i>SMARCB1</i> in established Malignant Rhabdoid cells lines 120</b>	
4.3.1	<i>SMARCB1</i> causes growth arrest and senescence cell formation in Rhabdoid cell lines .....	120
4.3.2	<i>SMARCB1</i> induces <i>CDKN2A</i> , <i>CDKN1A</i> and <i>CD44</i> over-expression in Malignant Rhabdoid Tumour cell lines .....	125
<b>4.4</b>	<b>Effects of 5-aza-2-deoxycytidine treatment on methylation status of Malignant Rhabdoid cells .....</b>	<b>133</b>

<b>4.5</b>	<b>SMARCA4 ATPse subunits acts as proto-oncogene in Malignant Rhabdoid Tumour .....</b>	<b>137</b>
<b>4.6</b>	<b>Discussion .....</b>	<b>143</b>
4.6.1	Characterisation of the effects of <i>SMARCB1</i> loss in Malignant Rhabdoid cells.....	143
4.6.2	Effects of <i>SMARCB1</i> loss on methylation in Malignant Rhabdoid cells	144
4.6.3	Tumourigenesis induced by <i>SMARCB1</i> loss is affected by aberrant activity of SMARCA4 ATPse subunit.....	145
<b>Chapter 5 Genomic and Methylation Analysis of Malignant Rhabdoid Tumours .....</b>		<b>147</b>
<b>5.1</b>	<b>Introduction .....</b>	<b>148</b>
<b>5.2</b>	<b>Aims .....</b>	<b>150</b>
<b>5.3</b>	<b>RNAseq analysis identifies Malignant Rhabdoid Tumours as a single biological entity containing subgroups .....</b>	<b>151</b>
5.3.1	Primary samples preparation.....	151
5.3.2	RNA-seq QC analysis .....	153
5.3.3	QC analysis .....	156
5.3.3.1	<i>Unsupervised analysis</i> .....	156
5.3.3.2	<i>Supervised analysis</i> .....	161
5.3.4	Cluster analysis identifies two major subgroup in the Rhabdoid cohort mainly related to location of the malignacies.....	169
5.3.4.1	<i>Unsupervised analysis</i> .....	169

	<i>Supervised analysis</i> .....	174
	<i>RNA-seq metagene projection onto existing primary Malignant Rhabdoid Tumours microarray data sets</i> .....	180
5.3.5	DEXSeq analysis: differential exon usage .....	183
5.3.5.1	<i>Detection of differentially used exons between Malignant Rhabdoid Tumours and Medulloblastoma primary tumours</i> .....	186
5.3.5.2	<i>Detection of differentially used exons between Extra Cranial Rhabdoids tumours and Atypical Teratoid Rhabdoid Tumours.</i> .....	191
<b>5.4</b>	<b>450K methylation array confirm different biological group in CNS tumours</b> .....	<b>197</b>
5.4.1	<i>Primary samples preparation and sequencing</i> .....	197
5.4.2	QC analysis .....	197
5.4.3	Global patterns of DNA methylation in Rhabdoid Tumours .....	199
5.4.4	Cluster analysis identifies 5 distinct DNA methylation subgroups.....	200
5.4.4.1	<i>Unsupervised analysis</i> .....	200
5.4.4.2	<i>Supervised analysis</i> .....	204
5.4.5	Rhabdoid Tumours present different DNA methylation subgroups mainly related to location of malignancy .....	210
5.4.5.1	<i>Unsupervised analysis</i> .....	210
5.4.5.2	<i>Supervised analysis</i> .....	214
<b>5.5</b>	<b>Discussion</b> .....	<b>221</b>



5.5.1	Malignant Rhabdoid Tumours are a unique biological group, with distinct methylation and expression pattern .....	221
5.5.2	Gene expression and methylation status reveals subgroups in Rhabdoid Malignant Tumours, mainly related to tumour location .....	223
<b>Chapter 6 Establishing the Role of <i>SMARCB1</i> loss in Malignant Rhabdoid Tumours.....</b>		<b>225</b>
<b>6.1</b>	<b>Introduction .....</b>	<b>226</b>
<b>6.2</b>	<b>Aims .....</b>	<b>227</b>
<b>6.3</b>	<b>Integrated genomic and epigenetic analysis identifies downstream target genes dependent on <i>SMARCB1</i> .....</b>	<b>228</b>
6.3.1	RNA-seq analysis.....	228
6.3.1.1	<i>Primary samples preparation.....</i>	228
6.3.1.2	<i>RNA-seq QC analysis .....</i>	228
6.3.1.3	<i>DESeq analysis.....</i>	231
	<i>Clustering analysis .....</i>	232
6.3.1.4	<i>DEXSeq analysis: differential exon usage .....</i>	239
6.3.1.5	<i>Detection of differentially used exon in gene <i>SMARCB1</i> dependent ..</i>	240
6.3.1.6	<i>Validation of the DEXseq analysis .....</i>	243
6.3.1.7	<i>Novel transcript discovery.....</i>	248
	<i>Transcript validation .....</i>	250
6.3.2	450K methylation analysis.....	257

6.3.2.1	<i>QC analysis</i> .....	257
6.3.2.2	<i>Global patterns of DNA methylation in Rhabdoid Tumours</i> .....	258
	<i>Cluster analysis</i> .....	259
<b>6.4</b>	<b>RNA-seq analysis identifies gene reactivation induced by 5-aza-2-deoxycytidine treatment in Malignant Rhabdoid cells</b> .....	<b>265</b>
6.4.1	RNA-seq analysis.....	265
6.4.1.1	<i>Primary samples preparation</i> .....	265
6.4.1.2	<i>RNA-seq QC analysis</i> .....	265
6.4.1.3	<i>DEseq analysis</i> .....	265
	<i>Clustering analysis</i> .....	267
	<i>DEXSeq analysis: differential exon usage</i> .....	272
6.4.1.4	<i>Detection of differentially used exon following by 5-aza-2-deoxycytidine treatment</i> .....	273
<b>6.5</b>	<b>Discussion</b> .....	<b>275</b>
6.5.1	Integrated genomic and epigenetic analysis identifies downstream target genes dependent by <i>SMARCB1</i> .....	275
6.5.2	RNA-seq reveals gene activation in Malignant Rhabdoid cell lines following 5-Aza-Deoxycytidine treatment.....	277
<b>Chapter 7</b>	<b>Comparison of Molecular Signatures from Malignant Rhabdoid Tumours and Cell line Models Identifies Mechanisms of Tumourigenesis</b> .....	<b>279</b>
<b>7.1</b>	<b>Introduction</b> .....	<b>280</b>

<b>7.2</b>	<b>Aims .....</b>	<b>282</b>
<b>7.3</b>	<b>Multiple platform data integration characterises the role of <i>SMARCB1</i> loss in epigenetic deregulation.....</b>	<b>283</b>
7.3.1	Cross-referencing 450K methylation data primary data and model data	283
7.3.1.1	<i>Validation of selected probes in Rhabdoid cell lines using bisulfite conversion.....</i>	<i>285</i>
	<i>SMARCB1 re-expression induces changes in ABR methylation status .....</i>	<i>286</i>
	<i>SMARCB1 re-expression induces changes in BHLHE4 methylation status ..</i>	<i>288</i>
	<i>SMARCB1 re-expression induces changes in GALNT2 methylation status...</i>	<i>290</i>
7.3.2	Gene expression cross-referencing in model data identifies <i>SMARCB1</i> role in epigenetic mechanism .....	292
7.3.2.1	<i>Cross-referencing expression data in model data reveals demethylating effects of SMARCB1 re-expression.....</i>	<i>292</i>
7.3.2.2	<i>Platform cross-referencing identifies genes directly demethylated by SMARCB1 .....</i>	<i>294</i>
<b>7.4</b>	<b>Pathway analysis.....</b>	<b>298</b>
7.4.1	Cross-referencing of multiple gene expression gene sets identifies gene signatures directly controlled by <i>SMARCB1</i> .....	298
7.4.2	Gene set enrichment analysis .....	305
7.4.2.1	<i>GSEA analysis of differentially regulated signalling pathways in primary Rhabdoid Tumours.....</i>	<i>306</i>

7.4.2.2	<i>GSEA analysis distinguishes differentially regulated signalling pathways in Extra Cranial Rhabdoid Tumours and Atypical Teratoid Rhabdoid Tumours .....</i>	313
7.4.2.3	<i>GSEA analysis of differentially regulated signalling pathways in Rhabdoid cell lines following SMARCB1 re-expression .....</i>	317
7.4.2.4	<i>GSEA analysis of differentially regulated signalling pathways in Rhabdoid cell lines treated with 5-aza-2-deoxycytidine.....</i>	323
7.4.3	<b>Ingenuity Pathway Analysis.....</b>	327
7.4.3.1	<i>Analysis of pathway deregulation in SMARCB1 re-expressing Rhabdoid Tumour cells .....</i>	327
7.4.3.2	<i>Analysis of pathway deregulation in Rhabdoid Tumours.....</i>	330
7.4.3.3	<i>Analysis of pathway deregulation in 5-aza-CdR treated Rhabdoid Tumour cells .....</i>	332
7.4.3.4	<i>Custom pathway analysis .....</i>	334
<b>7.5</b>	<b>Conclusion .....</b>	<b>340</b>
7.5.1	<b>Role of SMARCB1 in methylation modulation.....</b>	<b>340</b>
7.5.2	<b>Pathway analysis .....</b>	<b>341</b>
7.5.2.1	<i>Gene expression cross-referencing identifies signature of Rhabdoid phenotype.....</i>	341
7.5.2.2	<i>GSEA analysis identifies deregulated pathways responsible for tumourigenesis in Rhabdoid Tumours .....</i>	341
7.5.2.3	<i>Ingenuity Pathway analysis.....</i>	343

<b>Chapter 8 Conclusion.....</b>	<b>344</b>
<b>8.1 Introduction .....</b>	<b>345</b>
<b>8.2 DNA methylation and gene expression profiling of Rhabdoid Tumours characterise a unique a coherent tumour type when compared with other embryonal and neuronal malignancies.....</b>	<b>347</b>
8.2.1 Patterns of methylation and gene expression may have utility for disease sub-classification .....	348
<b>8.3 Epigenetic effect of <i>SMARCB1</i> deletion if Rhabdoid Tumours.....</b>	<b>350</b>
<b>8.4 Integrative analysis of model and primary data indicate activation of a variety of pathways dependent on <i>SMARCB1</i> deletion .....</b>	<b>352</b>
<b>8.5 Limitations .....</b>	<b>354</b>
<b>8.6 Future work .....</b>	<b>356</b>
8.6.1 Implication of gene deregulation in Rhabdoid Tumourigenesis .....	356
8.6.2 Characterisation of pathways .....	356
8.6.3 Characterisation of <i>SMARCA4</i> D/N effect in gene expression and methylation status .....	357
8.6.4 Single cell sequencing.....	357
<b>8.7 Final Conclusion .....</b>	<b>358</b>
<b>Chapter 9 References .....</b>	<b>359</b>
<b>Appendix I.....</b>	<b>443</b>
<b>450K methylation array.....</b>	<b>444</b>

De novo analysis .....	444
DESeq analysis: .....	444
NMF gene signature .....	444
DEXseq	445
GESEA	446
Ingenuity .....	446
Publications.....	447
Systems Biology Approach to Identify Mechanism of Tumourigenesis Caused by Loss of <i>SMARCB1</i> in Malignant Rhabdoid Tumors .....	448
Characterising critical pathways in ATRT tumourigenesis: a genome-wide analysis of primary tumours and functional genomic models .....	450
Mechanism of tumourigenesis caused by loss of <i>SMARCB1</i> in Malignant Rhabdoid Tumors .....	452
Next-generation sequencing identifies the mechanism of tumourigenesis caused by loss of <i>SMARCB1</i> in Malignant Rhabdoid Tumours.....	455
Awards and achievement.....	457

## List of figures

<i>Figure 1.1 Immunohistochemistry of Atypical Teratoid Rhabdoid Tumour. ....</i>	<i>3</i>
<i>Figure 1.2 Magnetic resonance image of a patient with Atypical Teratoid Rhabdoid Tumour.....</i>	<i>4</i>
<i>Figure 1.3 Relative frequencies of the 6 most common malignant central nervous system (CNS) tumours.....</i>	<i>5</i>
<i>Figure 1.4 Kaplan-Meier curves illustrate overall survival (OS) for patients with Aypical Teratoid Rhabdoid Tumours.....</i>	<i>6</i>
<i>Figure 1.5 SNP array analysis of chromosome 22 copy number alterations. ....</i>	<i>7</i>
<i>Figure 1.6 Kaplan-Meier curves illustrate overall survival (OS) for patients characterised by SMARCB1 germline mutation. ....</i>	<i>8</i>
<i>Figure 1.7 Observed frequency of SMARCB1 germline mutation. ....</i>	<i>9</i>
<i>Figure 1.8 Kaplan–Meier plots of the estimated PFS rate for 19 study patients. ....</i>	<i>11</i>
<i>Figure 1.9 Embryonal tumour microscopy features.....</i>	<i>14</i>
<i>Figure 1.10 HES (A-C) and anti-SMARCB1 (B-D) Immunohistochemical staIning of SMARCB1 negative tumours non compatible with Rhabdoid morphology. ....</i>	<i>15</i>
<i>Figure 1.11 Age at diagnosis of patient with SMARCB1 negative malignancies. ....</i>	<i>16</i>
<i>Figure 1.12 The coding and structural mechanism of chromatin.....</i>	<i>17</i>
<i>Figure 1.13 Schematic representation of the mammalian DNA methylation mechanism. ....</i>	<i>19</i>
<i>Figure 1.14 Structure of a nucleosome: DNA wound around a histone core. ....</i>	<i>21</i>

<i>Figure 1.15 Most common histone modifications, including acetylation, methylation and phosphorylation at several amino acid residues.....</i>	<i>22</i>
<i>Figure 1.16 Model of histone modifications in normal human cells. ....</i>	<i>24</i>
<i>Figure 1.17 Mechanism of transcription activation driven by chromatin remodelling complexes. ....</i>	<i>26</i>
<i>Figure 1.18: Scheme of the bulge mechanism. The DNA is wrapped around the core histone protein octamer. ....</i>	<i>29</i>
<i>Figure 1.19 SWI/SNF complexes facilitate activation of gene expression by interaction with TATA-binding proteins.....</i>	<i>31</i>
<i>Figure 1.20: Structure of the SWI/SNF complex figure shows direct interaction between SMARCB1 and SMARCA4/BRM with tumour suppressor proteins (red), oncogenes (blue) and AKT. ....</i>	<i>33</i>
<i>Figure 1.21 Multiple mutations of the mammalian SWI/SNF subunits can occur in human cancers. ....</i>	<i>35</i>
<i>Figure 1.22 Survival of SMARCB1 target mice. ....</i>	<i>38</i>
<i>Figure 1.23 RNA workflow.....</i>	<i>44</i>
<i>Figure 1.24 ChIP-seq work flow.....</i>	<i>46</i>
<i>Figure 1.25 Overview of the Infinium Human Methylation 450 BeadChip method. ....</i>	<i>48</i>
<i>Figure 3.1 Overview of DNA sequencing using the Illumina platform.....</i>	<i>86</i>
<i>Figure 3.2 Overview of the primary data analysis.....</i>	<i>88</i>
<i>Figure 3.3 Illustration of the three overlaps resolution modes, used in the high-throughput sequencing analysis.....</i>	<i>90</i>



<i>Figure 3.4 Example of a contig assembled by the joining of many short reads .....</i>	<i>93</i>
<i>Figure 3.5 Example of fusion events between gene A and gene B. ....</i>	<i>95</i>
<i>Figure 3.6. Patterns observed for paired-end fusion analysis. ....</i>	<i>96</i>
<i>Figure 3.7 Infinium Human Methylation 450K BeadChip workflow. ....</i>	<i>99</i>
<i>Figure 3.8 Example of similarity of intensity distributions between Infinium I and II probes with the same number of underlying CpGs. ....</i>	<i>103</i>
<i>Figure 3.9 Overall beta value distribution produced by InfI and InfII results in an aberrant distribution. ....</i>	<i>103</i>
<i>Figure 3.10 Example of hierarchical clustering. ....</i>	<i>106</i>
<i>Figure 3.11 Type of distance between clusters calculation methods. ....</i>	<i>107</i>
<i>Figure 3.12 Difference of methods to assess dissimilarity or affinity between pairs of objects. ....</i>	<i>108</i>
<i>Figure 3.13 Example of PCA orthogonal transformation. ....</i>	<i>109</i>
<i>Figure 3.14 Example of NMF analysis of expression data sets composed by N genes, M samples and imposed k=2 .....</i>	<i>111</i>
<i>Figure 4.1 SMARCB1 induction affects gene expression in the G401 cell line (time course).....</i>	<i>121</i>
<i>Figure 4.2 Change in cell morphology after SMARCB1 induction in Rhabdoid cell line (7 days).....</i>	<i>122</i>
<i>Figure 4.3 Effect of SMARCB1 re-expression on survival (growth) in G401 and A204 Rhabdoid cells.....</i>	<i>123</i>

<i>Figure 4.4 Effect of SMARCB1 re-expression on survival (growth) in STA-WT1 and CHLA-266 Rhabdoid cells. ....</i>	<i>124</i>
<i>Figure 4.5 SMARCB1 re-expression induces changes in gene expression in G401 and A204. ....</i>	<i>126</i>
<i>Figure 4.6 SMARCB1 re-expression induces changes in gene expression in G401 and A204. ....</i>	<i>127</i>
<i>Figure 4.7 SMARCB1-dependent recruitment of SMARCA4 at the CDKN2A (p16) and CDKN1A (p21) locus after SMARCB1 re-expression in G401 and A204 cells. ....</i>	<i>129</i>
<i>Figure 4.8 Recruitment of SMARCB1 at the CDKN2A (p16) and CDKN2A (p21) locus after SMARCB1 re-expression in G401 cells. ....</i>	<i>130</i>
<i>Figure 4.9 Recruitment of H3k4me3 histone modification at the CDKN2A (p16) and CDKN2A locus after SMARCB1 re-expression in G401 and A204 cells. ....</i>	<i>131</i>
<i>Figure 4.10 Recruitment of H3k27me3 histone modification at the CDKN2A (p21) and CDKN2A locus following SMARCB1 re-expression in G401 and A204 cells. ....</i>	<i>132</i>
<i>Figure 4.11 RT-PCR for DAZL (A), RASSF1A (B), HTATIP2(C) and GAPDH (D) in Malignant Rhabdoid Tumour cells treated with 5-azaCdR. ....</i>	<i>134</i>
<i>Figure 4.12 5-azaCdR induces changes in CDKN2A and CDKN1A expression in Rhabdoid cells in a similar manner to that observed following SMARCB1 re-expression. ....</i>	<i>135</i>
<i>Figure 4.13 5-azaCdR induces change in CD44 expression gene expression in Rhabdoid cells in a similar manner to that observed in SMARCB1 re-expression. ....</i>	<i>136</i>
<i>Figure 4.14 Change in cell morphology after SMARCA4 induction in Rhabdoid cell line ....</i>	<i>138</i>

<i>Figure 4.15 Effect of SMARCA4 ATPse activity inhibition on survival (growth) in G401 and A204 Rhabdoid cells. ....</i>	<i>139</i>
<i>Figure 4.16 Effect of SMARCA4 expression on apoptosis in the G401 cell line. Expression of SMARCA4 D/N induces strong early apoptosis which compromises cell viability.....</i>	<i>140</i>
<i>Figure 4.17 Effect of SMARCA4 inhibition on apoptosis in the A204 cell line. Expression of SMARCA4 D/N induces strong early apoptosis which compromises cell viability.....</i>	<i>141</i>
<i>Figure 4.18 Expression of SMARCA4 induces change in expression of CDKN1A, but not CDKN2A and CD44 in G401 and A204 cells .....</i>	<i>142</i>
<i>Figure 5.1 Electropherogram of sample nmb845 with 8.3 RIN score.....</i>	<i>152</i>
<i>Figure 5.2 Per base sequence quality derived from sample nmb 845 from a FASTQC html report.....</i>	<i>153</i>
<i>Figure 5.3 Per base sequence content analysis of sample nmb843 from FASTQC html report.....</i>	<i>154</i>
<i>Figure 5.4 Per base GC content of sample mnb 843 from FASTQC html report.....</i>	<i>155</i>
<i>Figure 5.5 Sequence duplication levels plot of sample mnb483 from FASTQC html report.....</i>	<i>155</i>
<i>Figure 5.6 Unsupervised Principal Component Analysis of DEseq data set distinguishes two different biological groups .....</i>	<i>157</i>
<i>Figure 5.7 Visualisation of cophenetic correlation and dispersion measures to assess the stability of clustering associated with each rank k for the Rhabdoid and Medulloblastoma derived DEseq data. ....</i>	<i>158</i>

<i>Figure 5.8 Unsupervised NMF and hierarchical cluster analysis (K=4) on DEseq data set characterises four different metagene classes. ....</i>	<i>159</i>
<i>Figure 5.9 Heatmap of gene expression showing the 20 most significantly differentially regulated genes between Medulloblastoma and Rhabdoid Tumours as calculated by adjusted p-value from DESeq2.....</i>	<i>160</i>
<i>Figure 5.10 Gene expression dispersion estimation and model fitting of Rhabdoid and Medulloblastoma derived DEseq data. ....</i>	<i>162</i>
<i>Figure 5.11 Per gene dispersion estimates from Rhabdoid and Medulloblastoma derived DEseq data. Differentially expressed genes Log2fold versus mean expression over all samples.....</i>	<i>162</i>
<i>Figure 5.12 Supervised principal component analysis of 157 genes selected as significantly differentially expressed between Rhabdoid Tumours and Medulloblastoma distinguishes two different biological groups. ....</i>	<i>163</i>
<i>Figure 5.13 Supervised heatmap of gene expression in Medulloblastoma and Rhabdoid Tumours.....</i>	<i>164</i>
<i>Figure 5.14 Relative difference in expression of 1-12 of most deregulated genes between Rhabdoid Tumours and Medulloblastomas.....</i>	<i>167</i>
<i>Figure 5.15 Relative difference in expression of 13-20 most deregulated genes between Rhabdoid Tumours and Medulloblastomas.....</i>	<i>168</i>
<i>Figure 5.16 Unsupervised principal component analysis on DEseq data set identifies two major subgroups within the Rhabdoid cohort. ....</i>	<i>170</i>
<i>Figure 5.17 Visualisation of cophenetic correlation and dispersion measures to assess the stability of clustering associated with each rank k for the Rhabdoid derived DEseq data.....</i>	<i>171</i>

<i>Figure 5.18 Unsupervised NMF and hierarchical cluster analysis (K=2) on RNAseq data sets characterises two metagene classes mainly correlated with the location of the primary Rhabdoid Tumours .....</i>	<i>172</i>
<i>Figure 5.19 Heatmap of gene expression showing the 20 most highly expressed and differentially regulated genes between Extra Cranial Rhabdoid and Atypical Teratoid Rhabdoid Tumours of the brain. ....</i>	<i>173</i>
<i>Figure 5.20 Supervised principal component analysis on DEseq data set identifies two major subgroups in Rhabdoid cohort.....</i>	<i>174</i>
<i>Figure 5.21 Heatmap of gene expression showing the highly expressed and differentially regulated genes between Extra Cranial Rhabdoid and Atypical Teratoid Rhabdoid Tumours of the brain .....</i>	<i>175</i>
<i>Figure 5.22 Relative differences in expression of 1-12 most deregulated genes between Extra Cranial Rhabdoid Tumours and Atypical Teratoid Rhabdoid Tumours .....</i>	<i>178</i>
<i>Figure 5.23 Relative difference in expression of 13-20 most deregulated genes between Extra Cranial Rhabdoid Tumours and Atypical Teratoid Rhabdoid Tumours (ATRT). ....</i>	<i>179</i>
<i>Figure 5.24 Heatmaps of metagene projection of Malignant Rhabdoid Tumour primary samples. These heatmaps show the metagene expression levels for each sample. ....</i>	<i>181</i>
<i>Figure 5.25 Heatmap of metagene projection of Malignant Rhabdoid Tumour primary samples.....</i>	<i>182</i>
<i>Figure 5.26 Differential exon usage dispersion estimation and model fitting of Rhabdoid and Medulloblastoma derived DEXseq data.....</i>	<i>184</i>
<i>Figure 5.27 Per exonic part dispersion estimates of Rhabdoid and Medulloblastoma derived DEXseq data.....</i>	<i>184</i>
<i>Figure 5.28 Differential exon usage dispersion estimation and model fitting of Rhabdoid Tumour DEXseq data. ....</i>	<i>185</i>

<i>Figure 5.29 Differentially used exons Log2fold change versus mean expression over all samples.....</i>	<i>185</i>
<i>Figure 5.30 HES6 shows differential exon usage in Rhabdoid Tumours compared with Medulloblastoma sample .....</i>	<i>188</i>
<i>Figure 5.31 TUBA1C shows differential exon usage in Medulloblastoma tumours compared with Rhabdoid sample .....</i>	<i>189</i>
<i>Figure 5.32 Differential exon usage of RPS27L in Rhabdoid sample tumours .....</i>	<i>190</i>
<i>Figure 5.33 Differential exon usage of DCLK1 in Rhabdoid sample tumours.....</i>	<i>193</i>
<i>Figure 5.34 Differential exon usage of RNF220 in Rhabdoid sample tumours.....</i>	<i>194</i>
<i>Figure 5.35 Differential exon usage of HOPX in Rhabdoid Tumours .....</i>	<i>196</i>
<i>Figure 5.36 Control strip plot of 450K methylation data. ....</i>	<i>198</i>
<i>Figure 5.37 Overall density distributions of beta values and the density distributions in primary Rhabdoid cohort.....</i>	<i>199</i>
<i>Figure 5.38 Unsupervised principal component analysis of 450K methylation profiles of primary Medulloblastoma and Rhabdoid Tumours distinguishes two biological groups. ....</i>	<i>200</i>
<i>Figure 5.39 Cophenetic correlation and dispersion measures to assess the stability of clustering associated with each rank k for the methylation data set. ....</i>	<i>201</i>
<i>Figure 5.40 Unsupervised NMF and hierarchical cluster analysis (K=4) on Rhabdoid and Medulloblastoma derived 450K methylation data set.....</i>	<i>202</i>
<i>Figure 5.41 Heatmap of methylation showing the 100 most differentially methylated probes between Medulloblastoma and Rhabdoid Tumours. ....</i>	<i>203</i>

<i>Figure 5.42 Supervised principal component analysis of 450K methylation profiles of primary Medulloblastoma and Rhabdoid Tumours distinguishes two biological groups</i> .....	204
<i>Figure 5.43 Relative difference in methylation status of 12 of 20 most deregulated genes between Rhabdoid Tumours and Medulloblastomas</i> .....	207
<i>Figure 5.44 Relative difference in methylation status of 8 of 20 most deregulated genes between Rhabdoid Tumours (RT) and Medulloblastomas (MB)</i> .....	208
<i>Figure 5.45 Unsupervised principal component analysis on 450K methylation data set characterises two subgroups within Rhabdoid Tumours</i> .....	210
<i>Figure 5.46 Cophenetic correlation and dispersion measures to assess the stability of clustering associated with each rank k for the methylation data set. ....</i>	211
<i>Figure 5.47 Consensus unsupervised NMF and hierarchical cluster analysis k=2) on Rhabdoid Tumour derived 450K methylation data set</i> .....	212
<i>Figure 5.48 Heatmap of gene expression showing the 100 most differentially methylated probes between Atypical Teratoid Rhabdoid Tumours (ATRT) and Extra Cranial Rhabdoid Tumours (ECRT)</i> .....	213
<i>Figure 5.49 Supervised principal component analysis on Malignant Rhabdoid Tumour 450K methylation data set distinguishes two biological groups, related to the tumour location</i> .....	214
<i>Figure 5.50 Heatmap of gene expression showing the most differentially methylated probes in Rhabdoid Tumours in supervised fashion. ....</i>	215
<i>Figure 5.51 Relative difference in methylation status of 12 of the most deregulated genes between Atypical Teratoid Rhabdoid Tumours and Extra Cranial Rhabdoid Tumour</i> .....	218

<i>Figure 5.52 Relative difference in methylation status 8 of 20 most deregulated genes between Atypical Teratoid Rhabdoid Tumours and Extra Cranial Rhabdoid Tumours</i>	219
<i>Figure 6.1 Per base sequence quality derived from sample G401 SMARCB1+ from FASTQC html report.</i>	229
<i>Figure 6.2 Per base GC content (B) of sample G401 SMARCB1+ from FASTQC html report.</i>	229
<i>Figure 6.3 Per base sequence content analysis (A) and Per base GC content (B) of sample G401 SMARCB1+ from FASTQC html report.</i>	230
<i>Figure 6.4 Sequence duplication levels plot of sample G401 SMARCB1+ from FASTQC html report.</i>	230
<i>Figure 6.5 Gene expression dispersion estimation and model fitting of SMARCB1 re-expressing cells and control infectants</i>	231
<i>Figure 6.6 Testing for differential expression between SMARCB1 re-expressing cells and control infectants.</i>	232
<i>Figure 6.7 Supervised principal component analysis on Rhabdoid cell lines expression distinguishes two different biological groups.</i>	233
<i>Figure 6.8 Supervised NMF and hierarchical cluster analysis (K=2) on expression data from SMARCB1 re-expressing cells and control infectants</i>	234
<i>Figure 6.9 Heatmap of significantly differentially expressed genes in SMARCB1 re-expressing Rhabdoid cells and in vector control infected Rhabdoid cells</i>	235
<i>Figure 6.10 Relative difference in expression of 9 of 15 most deregulated genes between SMARCB1 re-expressing Rhabdoid cells and vector control infected Rhabdoid cells</i>	237



<i>Figure 6.11 Relative difference in expression of 6 of 15 most deregulated genes between SMARCB1 re-expressing Rhabdoid cells and vector control infected Rhabdoid cells .</i>	<i>238</i>
<i>Figure 6.12 Per exon usage dispersion estimation and model fitting and testing for differential exon usage of genes comparing SMARCB1 re-expressing cells and control infectants. ....</i>	<i>239</i>
<i>Figure 6.13 BTG4 shows differential exon usage in response to SMARCB1 re-expression .....</i>	<i>241</i>
<i>Figure 6.14 PGAM1 shows differential exon usage in response to SMARCB1 re-expression.....</i>	<i>242</i>
<i>Figure 6.15 MYLK shows differential exon usage in response to SMARCB1 re-expression.....</i>	<i>244</i>
<i>Figure 6.16 SMARCB1 re-expression induces differential usage of MYLK exon 4 in Malignant Rhabdoid cell lines.....</i>	<i>245</i>
<i>Figure 6.17 IL17RE differential exon usage in response to SMARCB1 re-expression .....</i>	<i>246</i>
<i>Figure 6.18 SMARCB1 re-expression induces differential exon usage of IL17RE exonic part 23, 28 and 33 in Malignant Rhabdoid cell lines. ....</i>	<i>247</i>
<i>Figure 6.19 IGV representation of raw read counts that correspond to XLOC_053533 and the difference between SMARCB1 re-expressing cells and the control infectants. ....</i>	<i>250</i>
<i>Figure 6.20 IGV representation of raw read counts that correspond to and XLOC_032441 and the difference between SMARCB1 re-expressing cells and the control infectants.....</i>	<i>251</i>
<i>Figure 6.21 RT-PCR for XLOC_053533 exon boundary 1- 2 and exon boundary 3-4 .....</i>	<i>252</i>

<i>Figure 6.22 RT-PCR for XLOC_053533 exon boundary 4-5 .....</i>	<i>253</i>
<i>Figure 6.23 Sanger sequencing of XLOC_053533 validated the presence of exon boundary 1-2, exon boundary 3-4, and exon boundary 4-5 at the locus chr7:40165570-41173105. ....</i>	<i>254</i>
<i>Figure 6.24 RT-PCR for XLOC_032441 exon boundary 1-2, exon boundary 2-3 and exon boundary 4-5. ....</i>	<i>255</i>
<i>Figure 6.25 Sanger sequencing validated the presence of exon boundary 2-3 in the locus XLOC_032441 chr2:150959288-150975090. ....</i>	<i>256</i>
<i>Figure 6.26 Control strip plot of 450K methylation data for Rhabdoid cell line models. ....</i>	<i>257</i>
<i>Figure 6.27 Overall density distributions of beta values in Rhabdoid cell line models. ....</i>	<i>258</i>
<i>Figure 6.28 Supervised principal component analysis on SMARCB1 re-expressing cell lines vs control infectants. ....</i>	<i>260</i>
<i>Figure 6.29 Supervised consensus NMF and hierarchical cluster analysis (K=2) performed on methylation data from SMARCB1 re-expressing cells and control infectants .....</i>	<i>261</i>
<i>Figure 6.30 Supervised heatmap of differentially methylated probes in SMARCB1 re-expressing Rhabdoid cells and in vector control infected Rhabdoid cells. ....</i>	<i>262</i>
<i>Figure 6.31 Gene expression dispersion estimation and model fitting of 5-azaCdR treated cells .....</i>	<i>266</i>
<i>Figure 6.32 Gene expression dispersion estimation and model fitting of control untreated cell lines. ....</i>	<i>267</i>

<i>Figure 6.33 Testing for differential expression between of 5-azaCdR treated cells vs untreated cells .....</i>	<i>267</i>
<i>Figure 6.34 Supervised principal component analysis of gene expression profiles from 5-azaCdR treated Rhabdoid cell lines and untreated Rhabdoid cell lines distinguishes two different biological groups .....</i>	<i>268</i>
<i>Figure 6.35 Supervised NMF and hierarchical cluster analysis (K=2) of 5-azaCdR treated cells and untreated cells.....</i>	<i>269</i>
<i>Figure 6.36 Supervised heatmap of selected significantly differentially expressed genes in 5-azaCdR treated cells and untreated cells.....</i>	<i>270</i>
<i>Figure 6.37 A) Per exon usage dispersion estimation and model fitting and testing for differential exon usage of genes comparing untreated cells and 5-azaCdR treated cells. ....</i>	<i>272</i>
<i>Figure 7.1 Transcription factor ChIP-Seq data for ABR derived from the Encode project .....</i>	<i>286</i>
<i>Figure 7.2 Sanger sequencing indicates demethylation of ABR in response to SMARCB1 re-expression in G401 cell lines.....</i>	<i>287</i>
<i>Figure 7.3 Transcription factor ChIP-Seq data for BHLHE4 derived from the Encode project.....</i>	<i>288</i>
<i>Figure 7.4 Sanger sequencing indicates demethylation of BHLHE4 in response to SMARCB1 re-expression in G401 cell line. ....</i>	<i>289</i>
<i>Figure 7.5 Transcription factor ChIP-Seq data for GALNT 2 derived from the Encode project.....</i>	<i>290</i>
<i>Figure 7.6 Sanger sequencing indicates demethylation of GALNT2 in response to SMARCB1 re-expression in G401 cell lines.....</i>	<i>291</i>

<i>Figure 7.7 Venn diagram of differentially expressed genes as result of SMARCB1 re-expression and 5-azaCdR treatment.....</i>	<i>292</i>
<i>Figure 7.8 SMARCB1 re-expression induces upregulation of GLI2 and GLI3 as a result of demethylation and/or SMARCB1 re-expression. ....</i>	<i>296</i>
<i>Figure 7.9 SMARCB1 re-expression induces upregulation of JARID2 and NTN4, as a result of demethylation and/or SMARCB1 re-expression.....</i>	<i>297</i>
<i>Figure 7.10 Cross-referencing of two metagene signatures from Malignant Rhabdoid Tumours Primary tumours vs Medulloblastomas and cell model data re-expressing SMARCB1 and control infectants reveals common significantly deregulated genes <math>p&lt;0.01</math> (red). ....</i>	<i>299</i>
<i>Figure 7.11 Cross-referencing of metagene signatures from Rhabdoid cells re-expressing SMARCB1 and from Rhabdoid cells treated with 5-azaCdR reveals a common gene signature in the data set .....</i>	<i>301</i>
<i>Figure 7.12 Venn diagram of Methylation-dependent Rhabdoid gene signatures common between SMARCB1 re-expressing cells and Medulloblastomas and between SMARCB1 re-expressing cells and 5-azaCdR treated cells. ....</i>	<i>303</i>
<i>Figure 7.13 GSEA enrichment plots of gene sets: Ren alveolar rhabdomyosarcomas, Wong embryonic stem cells core and Reactome p53 independent G1/S DNA damage checkpoint pathways genes in Malignant Rhabdoid Tumour vs Medulloblastomas....</i>	<i>308</i>
<i>Figure 7.14 GSEA enrichment plot of Asgharzadeh neuroblastoma poor survival, Le neuronal differentiation, Kim all disorders oligodendrocyte number corr DN in Malignant Rhabdoid Tumour vs Medulloblastomas .....</i>	<i>309</i>
<i>Figure 7.15 GSEA enrichment plot of YAP1, GLI1 upregulated pathways genes and CAHOY NEURONAL downregulated pathways in Malignant Rhabdoid Tumour vs Medulloblastomas .....</i>	<i>311</i>

<i>Figure 7.16 GSEA enrichment plot of two gene sets Wong Embryonic stem cell core gene and the Lee PTCH1 and SUFU in Atypical Teratoid Rhabdoid Tumours vs Extra Cranial Rhabdoid Tumour. ....</i>	<i>314</i>
<i>Figure 7.17 GSEA enrichment plot of KRAS pathways and Cahoy neuronal pathway in Atypical Teratoid Rhabdoid Tumours vs Extra Cranial Rhabdoid Tumour.....</i>	<i>316</i>
<i>Figure 7.18 GSEA enrichment plots of significantly enriched gene sets including genes with promoters bound by E2F4, Oligodendrocyte differentiation pathway, and Reactome cell cycle annotation. ....</i>	<i>318</i>
<i>Figure 7.19 GSEA enrichment plot of significantly negatively enriched gene sets upregulated including Alveolar Rhabdomyosarcomas down-regulated genes, TBFB1 target pathways and PDEF target gene sets. ....</i>	<i>319</i>
<i>Figure 7.20 GSEA enrichment plot of genesets upregulated in response to RB and RBL1, and genesets upregulated in response to E2F3 and E2F1 .....</i>	<i>321</i>
<i>Figure 7.21 GSEA enrichment plot of gene sets related to HOXA9, TGFB1 pathways. These gene sets are significantly negatively enriched after SMARCB1 re-expression vs the control infectants.....</i>	<i>322</i>
<i>Figure 7.22 GSEA enrichment plots of selected significantly enriched genesets describing targets of HDAC1 and HDAC2, down regulated in response to TGFbeta activation and EZH2 targets. ....</i>	<i>324</i>
<i>Figure 7.23 GSEA enrichment plot of selected significantly enriched genesets including targets of HOXA9, NOTCH and SMARCB1 Knockout. ....</i>	<i>326</i>
<i>Figure 7.24 Important biological functions associated with genes that defined SMARCB1 re-expressing cell lines.....</i>	<i>328</i>
<i>Figure 7.25 Selected canonical pathways associated with genes that defined SMARCB1 re-expressing cell lines.....</i>	<i>329</i>

<i>Figure 7.26 Important biological functions associated with genes whose expression defines primary Rhabdoid Tumours.....</i>	<i>330</i>
<i>Figure 7.27 Selected canonical pathways associated with genes whose expression defines Rhabdoid Primary tumours. ....</i>	<i>331</i>
<i>Figure 7.28 Selected significant biological functions associated with 5-azaCdR treatment in Rhabdoid Tumours.....</i>	<i>332</i>
<i>Figure 7.29 Selected canonical pathways associated with 5-azaCdR treatment in Rhabdoid Tumours.....</i>	<i>333</i>
<i>Figure 7.30 Heatmap of the top 20 common canonical pathways identified by comparative analysis.....</i>	<i>335</i>
<i>Figure 7.31 Heatmap of the top 20 upstream effectors identified by comparative analysis. Intensity represents standard deviations from the mean.....</i>	<i>336</i>
<i>Figure 7.32 Modulation of the TGF-beta pathway is associated with Rhabdoid Tumourigenesis .....</i>	<i>337</i>
<i>Figure 7.33 Modulation of the TGF-beta pathway is associated with Rhabdoid Tumourigenesis .....</i>	<i>338</i>
<i>Figure 7.34 Modulation of the TGF-beta pathway is associated with Rhabdoid Tumourigenesis .....</i>	<i>339</i>

## List of tables

<i>Table 2.1 Primary Atypical Teratoid Rhabdoid Tumours cohort clinical data.</i>	<i>53</i>
<i>Table 2.2 Primary Extra Cranial Rhabdoid Tumours cohort clinical data.</i>	<i>54</i>
<i>Table 2.3 Media used for culturing cell lines.</i>	<i>56</i>
<i>Table 2.4 Buffers used for Western Blot</i>	<i>58</i>
<i>Table 2.5 Antibodies used for Western Blot.</i>	<i>59</i>
<i>Table 2.6 q-PCR oligonucleotides used to evaluate differential expression.</i>	<i>61</i>
<i>Table 2.7 General PCR reaction and conditions</i>	<i>62</i>
<i>Table 2.8 Buffer used for bacterial culture and composition</i>	<i>66</i>
<i>Table 2.9 Restriction endonuclease used for DNA manipulation</i>	<i>68</i>
<i>Table 2.10 Oligonucleotides used for screening of bacterial colonies and DNA manipulation.</i>	<i>69</i>
<i>Table 2.11 General PCR reaction conditions for each polymerase.</i>	<i>70</i>
<i>Table 2.12 Conditions used for blunt and cohesive end ligations.</i>	<i>72</i>
<i>Table 2.13 Oligonucleotides used for sequencing recombinant plasmid</i>	<i>74</i>
<i>Table 2.14 Antibodies used for Chromatin Immunoprecipitation</i>	<i>76</i>
<i>Table 2.15 Oligonucleotide used for sequencing bisulfite-converted DNA.</i>	<i>77</i>
<i>Table 2.16 General PCR reaction conditions for each polymerase used to amplify DNA intended to sequencing.</i>	<i>78</i>
<i>Table 2.17 Oligonucleotides used for sequencing bisulfite-converted DNA.</i>	<i>80</i>

<i>Table 2.18 Oligonucleotides used for sequencing bisulfite-converted DNA.....</i>	<i>81</i>
<i>Table 3.1. Qphred quality score values and accuracy .....</i>	<i>88</i>
<i>Table 5.1 Differentially expressed genes in Rhabdoid vs Medulloblastoma primary tumours.....</i>	<i>166</i>
<i>Table 5.2 Differentially expressed gene in Extra Cranial Rhabdoid vs Atypical Teratoid Rhabdoid Tumours. ....</i>	<i>177</i>
<i>Table 5.3 Genes with differentially used exons in Rhabdoid vs Medulloblastoma primary tumours.....</i>	<i>187</i>
<i>Table 5.4 Genes with differentially used exon in Rhabdoid primary tumours. ....</i>	<i>192</i>
<i>Table 5.5 Global methylation distribution among Malignant Rhabdoid primary tumours and Medulloblastomas test cohorts.....</i>	<i>199</i>
<i>Table 5.6 Selected significantly differentially methylated probes in Rhabdoid vs Medulloblastoma primary tumour .....</i>	<i>206</i>
<i>Table 5.7 Differentially methylated probes in Extra Cranial vs Atypical Teratoid Rhabdoid Tumours .....</i>	<i>217</i>
<i>Table 6.1 Selected differentially expressed genes in SMARCB1 re-expressing Rhabdoid cells vs vector control infected Rhabdoid cells. ....</i>	<i>236</i>
<i>Table 6.2 Genes with differentially used exons in SMARCB1 re-expressing Rhabdoid cells vs control infected cells.....</i>	<i>240</i>
<i>Table 6.3 List of significant possible novel transcript individualised in de novo transcriptome analysis. ....</i>	<i>248</i>
<i>Table 6.4 Analysis of significant transcript characterised by de novo transcriptome analysis.....</i>	<i>249</i>



<i>Table 6.5 Global methylation distribution among the cell lines .....</i>	<i>259</i>
<i>Table 6.6 Selected differentially methylated CpGs in response to SMARCB1 re-expression.....</i>	<i>263</i>
<i>Table 6.7 Differentially expressed gene SMARCB1 re-expressing Rhabdoid cells vs vector control infected Rhabdoid cells.....</i>	<i>271</i>
<i>Table 6.8 Genes with differentially used exons in 5-azaCdR treated Rhabdoid cells and untreated cells. ....</i>	<i>274</i>
<i>Table 7.1 List of top 20 genes and their associated CpG dinucleotides resulting in significant hypomethylation following SMARCB1 re-expression as a result of cross-filtering of primary and model data. ....</i>	<i>284</i>
<i>Table 7.2 List of the common differentially expressed genes resulting from SMARCB1 re-expression and 5-azaCdR data cross-referencing. ....</i>	<i>293</i>
<i>Table 7.3 Top 20 genes that define the Non- Rhabdoid phenotype resulting from cross-referencing Medulloblastomas metagene content and SMARCB1 re-expressing cell lines metagene content.....</i>	<i>300</i>
<i>Table 7.4 Top 20 genes that defined the Non-Rhabdoid phenotype resulting from cross-referencing 5-azaCdR treated cells metagene content and SMARCB1 re-expressing cell lines metagene content. ....</i>	<i>302</i>
<i>Table 7.5 Non Rhabdoid gene signatures common between SMARCB1 re-expressing cells and Medulloblastomas and between SMARCB1 re-expressing cells and 5-azaCdR treated cells. ....</i>	<i>304</i>
<i>Table 7.6 Selected C2 GSEA pathways upregulated and downregulated in primary Malignant Rhabdoid Tumours. ....</i>	<i>307</i>
<i>Table 7.7 3 selected C6 GSEA pathways significantly upregulated and downregulated in Rhabdoid Tumours when compared with Medulloblastomas. ....</i>	<i>310</i>

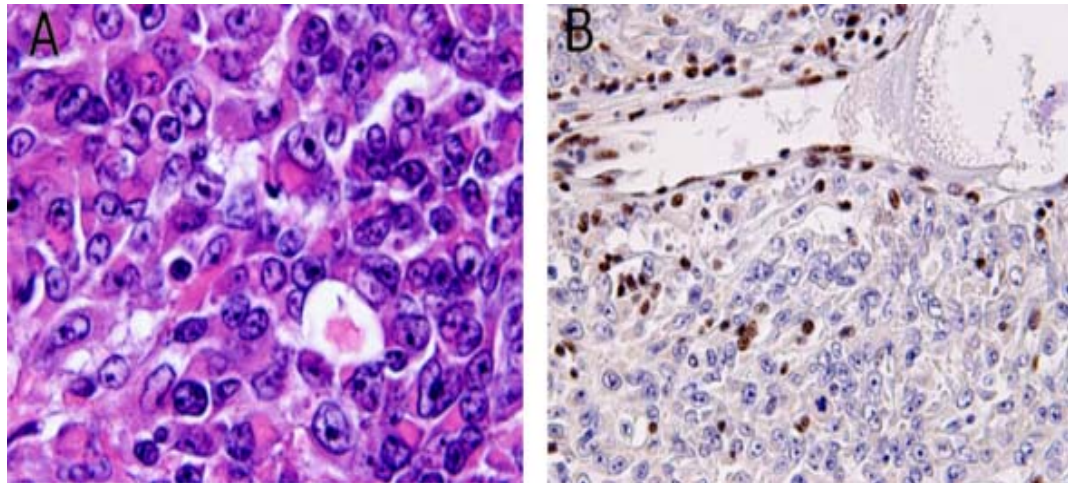
<i>Table 7.8 Two C2 library GSEA pathways upregulated and downregulated in Atypical Teratoid Rhabdoid Tumours when compared with Extra Cranial Rhabdoid Tumour. .</i>	<i>313</i>
<i>Table 7.9 C6 GSEA pathways upregulated and downregulated in Atypical Teratoid Rhabdoid Tumours when compared with Extra Cranial Rhabdoid Tumour.....</i>	<i>315</i>
<i>Table 7.10 Selected C2 GSEA pathways significantly enriched and negatively enriched in empty vector controls compared to SMARCB1 re-expressing cell lines .....</i>	<i>317</i>
<i>Table 7.11 Selected MSigDB C6 GSEA pathways significantly enriched or negatively enriched in empty vector controls compared to SMARCB1 re-expressing cell lines ....</i>	<i>320</i>
<i>Table 7.12 Selected C2 GSEA gene sets significantly enriched after 5-azaCdR treatment in Rhabdoid cell lines.....</i>	<i>323</i>
<i>Table 7.13 Selected MSigDB c6 significantly enriched and negatively enriched gene sets following 5-azaCdR treatment in Rhabdoid cell line .....</i>	<i>325</i>

# **Chapter 1      Introduction**

## **1.1 Atypical Teratoid Rhabdoid Tumours (ATRT)**

Atypical Teratoid Rhabdoid Tumour (ATRT) is an aggressive malignancy of the central nervous system (CNS). Central nervous system tumours represent 20% of all childhood neoplasms and are the most common cause of death in this age group (Peris-Bonet, Martinez-Garcia et al. 2006). Just 1- 2% of all paediatric brain tumours are Atypical Teratoid Rhabdoid Tumour., however they account for ~10% of all CNS tumours in infants with an incidence rate of 1.38 per 1,000,000 person-years in children (Woehrer, Slavic et al. 2010). New studies underline that in infancy the Atypical Teratoid Rhabdoid Tumour occur with a frequency similar to Medulloblastoma and CNS Primitive Neuroectodermal Tumour PNET (17.3%, 16.0%, and 13.3%, respectively) (Woehrer, Slavic et al. 2010).

Despite the fact that Atypical Teratoid Rhabdoid Tumours were first described as a Malignant Rhabdoid tumours (MRT) of the central nervous system in 1987, Atypical Teratoid Rhabdoid Tumours have been recognized as a unique class of primarily brain and spinal tumours in 1996 (Rorke, Packer et al. 1996). Atypical Teratoid Rhabdoid Tumour are characterised by a unique combination of Rhabdoid cells, primitive neuroectodermal cells and tumour cells with mesenchymal, neuronal or glial differentiation (Figure 1.1).



*Figure 1.1 Immunohistochemistry of Atypical Teratoid Rhabdoid Tumour: A) negative SMARCB1 staining. B) SMARCB1 positive staining in endothelial and inflammatory cells (Sigauke, Rakheja et al. 2006).*

On average 50% of these tumours reside in the cerebellum (Ronald 2005) or cerebellopontine angle (Kufe, Pollock et al. 2003), often with contiguous brain stem involvement. About 40% of tumours develop as a mass in the cerebral or suprasellar supratentorial regions (Figure 1.2). Despite a well characterised biology, Atypical Teratoid Rhabdoid Tumours can be misdiagnosed often as Medulloblastoma, primitive neuroectodermal tumour (PNET), choroid plexus carcinoma or germ cell tumour, since they develop in identical areas of the CNS (Biegel 1997, Bikowska, Grajkowska et al. 2011). The discovery of the role of *SMARCB1* gene as a tumour suppressor gene in Atypical Teratoid Rhabdoid Tumours (Versteeg, Sevenet et al. 1998, Sevenet, Sheridan et al. 1999, Biegel, Kalpana et al. 2002) has led to a more accurate diagnosis of the pathology. Nowadays screening for biallelic inactivation of the *SMARCB1* gene is often used as an adjunct to histology in the differential diagnosis of Rhabdoid Tumours (Haberler, Laggner et al. 2006).



*Figure 1.2 Magnetic resonance image of a patient with Atypical Teratoid Rhabdoid Tumour. This sagittal-plane magnetic resonance image of the brain shows a mass in the posterior fossa (Parwani, Stelow et al. 2005)*

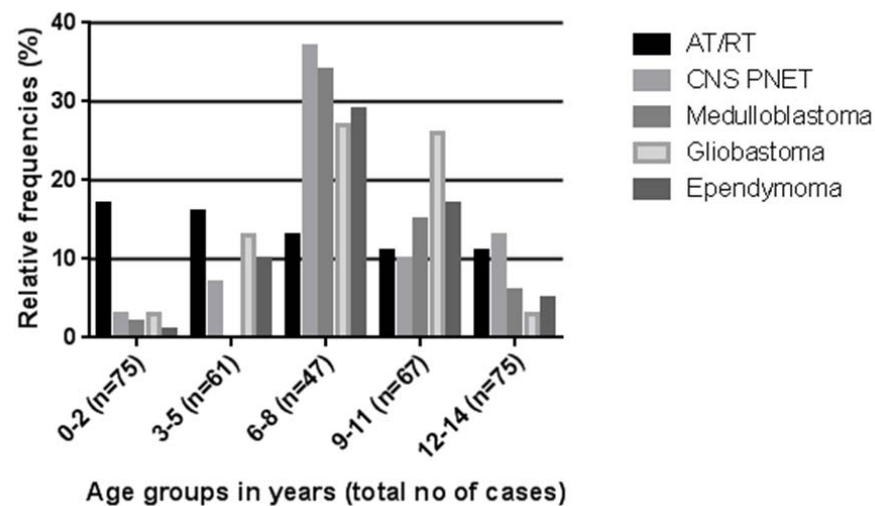
### **1.1.1 Epidemiology**

Atypical Teratoid Rhabdoid Tumour's incidence in childhood is still an approximate evaluation: often the diagnosis is inaccurate due to a lack of characteristic location and morphology and absence of systematic diagnostic evaluative protocols. However the discovery of changes in chromosome 22 and mutation in *SMARCB1* are leading to a better diagnosis, even retrospectively (Reinhard, Reinert et al. 2008)

Recent population-based study by the Austrain Brain tumour registry (Woehrer, Slavec et al. 2010) shows Atypical Teratoid Rhabdoid Tumours are more common than previously reported. In fact this cohort reveals that almost 50% of Atypical Teratoid Rhabdoid Tumours were misdiagnosed and in particular before the introduction of a systematic analysis of SMARCB1 protein (Burger, Yu et al. 1998, Hilden, Meerbaum et al. 2004, Haberler, Laggner et al. 2006). Nowadays, Atypical Teratoid Rhabdoid Tumours are considered one of the six most malignant central nervous system entities in childhood (Woehrer, Slavec et al. 2010). Atypical Teratoid Rhabdoid Tumours present a peak

incidence from birth to three years of age, with a median age of 1.44. In this age class Atypical Teratoid Rhabdoid Tumours have an incidence comparable with other CNS tumours (Woehrer, Slavc et al. 2010), indicating these malignancies are common especially in infants (Figure 1.3).

Atypical Teratoid Rhabdoid Tumours appear to be more aggressive than other CNS tumours, in fact the overall survival is 10 months from diagnosis and metastatic patients present a significantly shorter overall survival (Figure 1.4). A younger age at diagnosis provides a worse prognosis, although the overall survival is not affected by age (Buscariollo, Park et al. 2012).



*Figure 1.3 Relative frequencies of the 6 most common malignant central nervous system (CNS) tumours. PNET = Primitive Neuroectodermal Tumour ; patients are divided by age into five categories. In comparison to the other malignancies Atypical Teratoid Rhabdoid Tumours presents a peak incidence between 0 and 3 years age group (Woehrer, Slavc et al. 2010).*

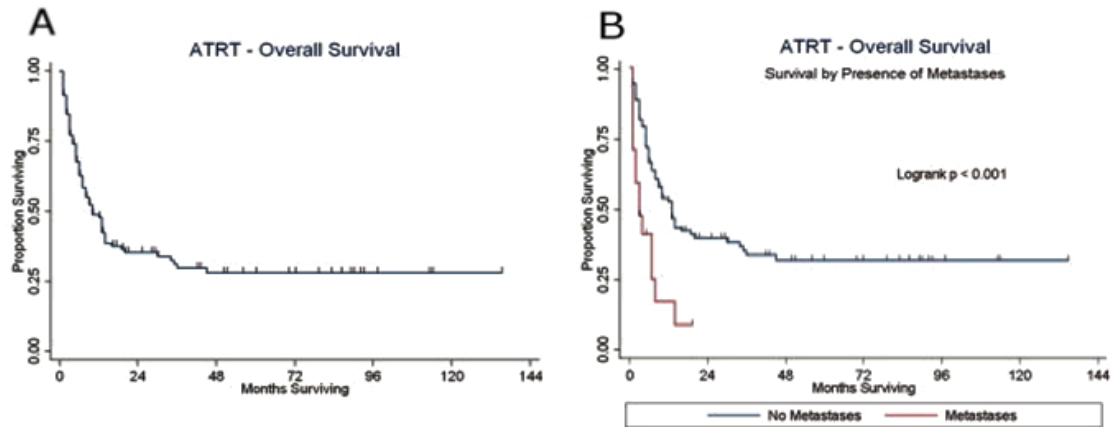


Figure 1.4 Kaplan-Meier curves illustrate overall survival (OS) for patients with Atypical Teratoid Rhabdoid Tumours. (A) The median OS for entire group was 10 months ( $n = 144$ ). (B) The median OS for patients who presented with metastases was 3 months; The median OS for patients without metastases was 13 months ( $P < .001$ ) (Buscariollo, Park et al. 2012).

### 1.1.2 Genetic epidemiology

In most Atypical Teratoid Rhabdoid Tumours deletions or abnormalities are detected in chromosome 22q11, implying the existence of a tumour suppressor gene at this locus. In fact Atypical Teratoid Rhabdoid Tumours (~50%) often present a loss of one copy of the entire chromosome 22 (monosomy 22) or a deletion or translocation in the band 22q11.2 (Versteeg, Sevenet et al. 1998). Moreover, Extra Cranial malignant Rhabdoid tumour (ECRT) show partial or complete deletions of chromosome 22 suggesting a common origin with Atypical Teratoid Rhabdoid Tumours, despite the differences in anatomic location.

Mutations have also been identified throughout the coding sequence, frequently occurring in exons 2 and 9, followed by 5, 4 and 7. Mutations in exons 2, 6 and 7 were commonly reported in kidney tumours (Eaton, Tooke et al. 2011), while exons 5 and 9 are frequently mutated in CNS tumours (Biegel, Tan et al. 2002). Single nucleotide polymorphism (SNP) oligonucleotide array analysis demonstrated that 98% of Rhabdoid Tumours (Atypical Teratoid Rhabdoid Tumours, renal and extra renal





deletion (Bourdeaut, Lequin et al. 2011, Eaton, Tooke et al. 2011). Presence of germline mutation was also described as a factor denoting poor prognosis (Tomlinson, Breslow et al. 2005, Kordes, Gesk et al. 2010, Bourdeaut, Lequin et al. 2011) (Figure 1.6).

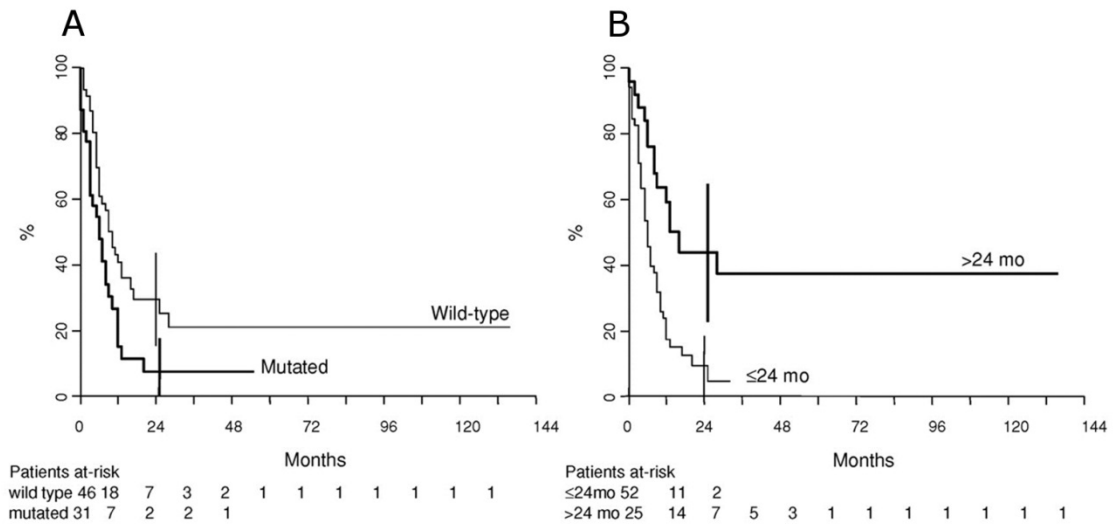
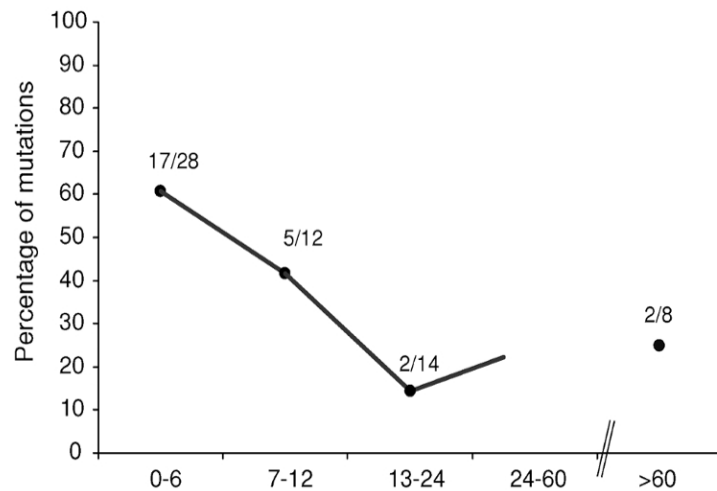


Figure 1.6 Kaplan-Meier curves illustrate overall survival (OS) for patients characterised by SMARCB1 germline mutation. A) Overall survival of Atypical Teratoid Rhabdoid Tumours patients with SMARCB1 germline mutation versus patients with normal constitutional SMARCB1 gene ( $p \leq 0.02$ ); B) Overall survival at tumour onset ( $p \leq 0.0004$ ) (Bourdeaut, Lequin et al. 2011).

The majority of mutations are acquired, but *de novo* mutations have been identified in the germline that are accompanied by somatic loss of the remaining allele (Biegel et al., 1999; 2002) and more frequent in patients younger than 2 years (Figure 1.7) (Bourdeaut, Lequin et al. 2011).



*Figure 1.7 Observed frequency of SMARCB1 germline mutation. Patients are divided by age into five categories. 60% of the Rhabdoid Tumours occurring before 6 months of age are correlated with SMARCB1 germline mutation (Bourdeaut, Lequin et al. 2011).*

### 1.1.3 Treatment

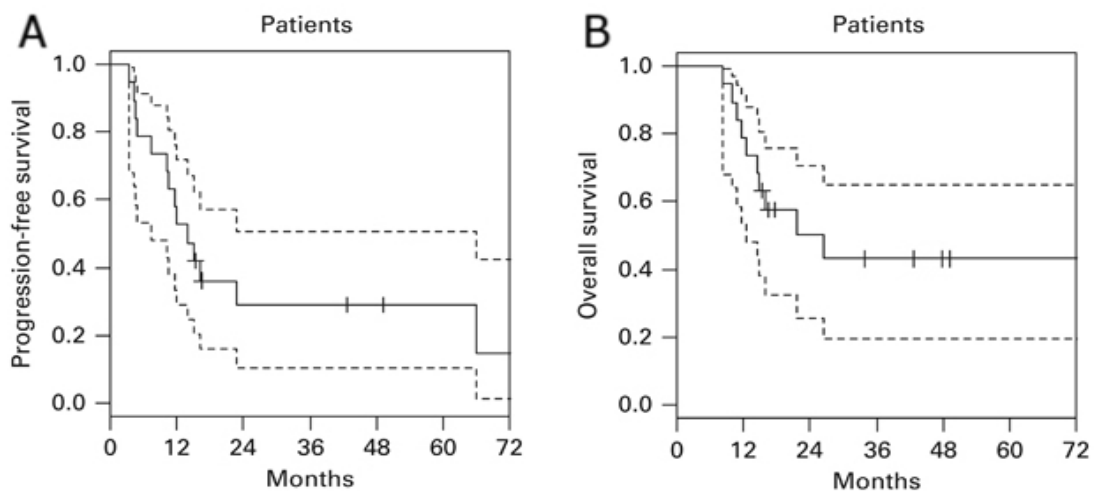
The most frequently used treatment for these neoplasms consists of intensive multi-modal therapy with surgery, chemotherapy and/or radiation therapy (Garre and Tekautz 2010). Initially surgical treatments are used to reduce the neoplastic mass, following by a supplementary therapy. The resection of the primary tumour mass is associated with improved survival, however gross total removal is possible in only 50% of cases and depends on the infiltration of the tumour (Bansal and Goel 2007).

In children younger than 2 years irradiation is typically avoided and high dose chemotherapy in conjunction with autologous stem cell transplantation are often used (Dunkel, Boyett et al. 1998). Recently an innovative treatment of cycles of carboplatin, doses of thiotepa and cycles of conventional induction chemotherapy resulted in 5 of 10 Atypical Teratoid Rhabdoid Tumour patients alive 42 months from diagnosis (Tekautz, Fuller et al. 2005). Radiotherapy is often given to children older than 3 years, due to

radio-toxicity which causes severe effects in developing brain. Radiotherapy prolongs the survival of older children but seems to be especially effective in the first stages of the regimen (Bansal and Goel 2007). Recent cohort studies demonstrate that radiotherapy had a strong impact on survival even in younger patients (<1 year and 1-3 years) (Buscariollo *et al*, 2012 (Benesch, Bartelheim *et al*. 2014)). Despite potentially severe side effects such as cognitive delay and endocrine dysfunction. Radiotherapy is still considered one of the most effective treatments for Atypical Teratoid Rhabdoid Tumours. Patients with localised mass can benefit from radiotherapy, since they can be treated with focal irradiation to minimise toxicities. In Europe treatment in patients younger than 18 months (Benesch, Bartelheim *et al*. 2014, Slavec, Chocholous *et al*. 2014) is more conservative and consists of a multistep therapy. In particular, in the United Kingdom after first surgery nine cycles of ifosfamide, carboplatin and etoposide (ICE), vincristine, cyclophosphamide, adriamycin (VCA) and doxorubicin are administered to patients between 0 and 12 months. After 7 months a second surgery is typically suggested and usually after 13 months of life the more aggressive high-dose-chemotherapy (HDCT) is adopted. HDCT is based on carboplatin and thiotepa given with a daily dose respectively of 500mg/m<sup>2</sup> and 300mg/m<sup>2</sup>. Radiotherapy is adopted in patients younger than 18 months just following the 9 cycles of chemotherapy. In patients older than 18 months radiotherapy is suggested to start as soon as possible and different strategies are taken on the basis of age and type of malignancies. In particular, in children with localised tumour mass irradiation with 36 Gy (1.8Gy daily, 5 fractions per week) is recommended. Metastatic patients aged between 18 months to 3 years are treated using craniospinal radiotherapy with 24.0 Gy (1.6 Gy daily), in children older than 3 years the axis dose is increased to 35.2 Gy (CCLG registry).

Retrospective analysis of the European Rhabdoid registry (EU-RHAB) suggests radiotherapy following gross total resection and HDCT could positively impact on the outcome of kids older than 18 months and with localised disease (Figure 1.8). However in non-irradiated patients (younger than 18 months) HDCT is only occasionally successful (Benesch, Bartelheim *et al*. 2014).

Despite the relatively positive outcome in older children, radiotherapy application remains controversial because of severe side effects. In particular, neurocognitive and vascular side effects have been observed especially in infants, even a long time after completion of the radiotherapy. The treatments used in the clinic for Atypical Teratoid Rhabdoid Tumours are often ineffective and with severe side effects that are the results of progressive attempts to increase survival in these patients. The importance of understanding the biological mechanisms that drive tumour formation in order to find an effective therapy is increasingly evident.



*Figure 1.8 Kaplan–Meier plots of the estimated PFS rate for 19 study patients. (b) Kaplan–Meier plots of the estimated OS rate for 19 study patients (Benesch, Bartelheim et al. 2014)*

#### **1.1.4 Extra Cranial Rhabdoid Tumours (ECRT)**

The Brain is not the only location in which Rhabdoid Tumours can be found. In fact Malignant Rhabdoid Tumours have been reported in other locations of the body, including the brain, liver, soft tissues, lung, skin, and heart. However Extra Cranial Rhabdoid Tumours (ECRT) have the same deletions of the *SMARCB1* gene indicating

that Atypical Teratoid Rhabdoid Tumours and ECRT are related entities, despite the different location. Moreover Rhabdoid Tumours at all locations display similar histological, clinical, and demographic features. Moreover, as well as Atypical Teratoid Rhabdoid Tumours, 15-30% of Extra Cranial Rhabdoid Tumours have been associated with germline mutation at chromosome 22q11.2 (Trobaugh-Lotrario, Finegold et al. 2011). Usually patients with *SMARCB1* germline mutation develop multiple Rhabdoid malignancies in first year of life and all have very poor prognoses regardless of the location of the tumour (Biegel 1999, Savla, Chen et al. 2000, Jackson, Shaikh et al. 2007).

Extra Cranial Rhabdoid Tumours can be further divided into renal and extra-renal malignancies. This sub-classification is due to the observation that Rhabdoid Tumours located in the kidneys tend to develop earlier and tend to progress in the brain (Tomlinson, Breslow et al. 2005). Patients with extra-renal Rhabdoid tend to be older and not show multiple metastasis or have a limited progression to the musculoskeletal system (Sultan, Qaddoumi et al. 2010).

Like Atypical Teratoid Rhabdoid Tumours, Extra Cranial ECRT are lethal and new treatments have indicated limited improvement on the survival. In particular in the UK just 31% of 106 the patients have 1-year survival from 1993-2010 (Brennan, Stiller et al. 2013).

#### **1.1.5 Rhabdoid predisposition syndrome in CNS tumours and tumours with *SMARCB1* germ line mutation**

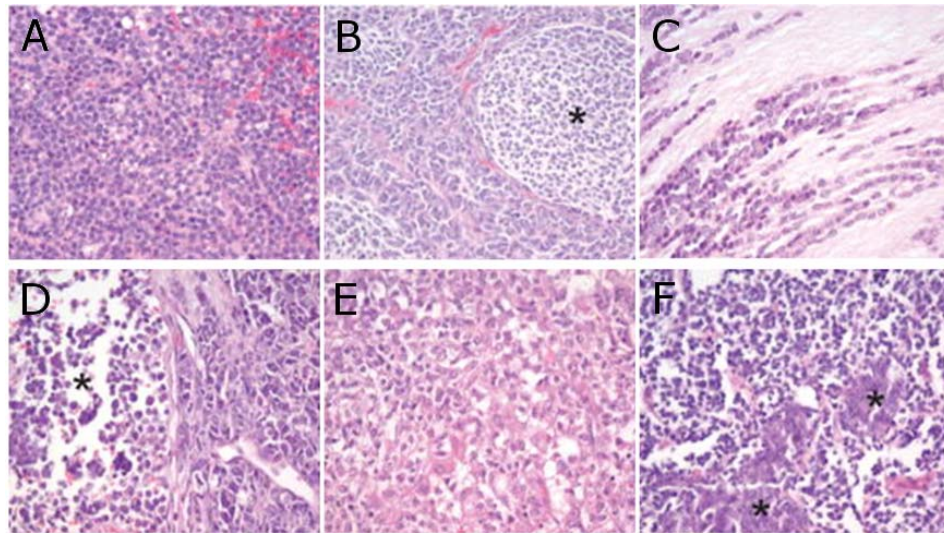
The new screening techniques have been helpful to highlight the hereditary nature of Rhabdoid tumour predisposition. Systematic investigation revealed that a significant number (30%) of patients harboured a germline *SMARCB1* mutation, regardless of the location of the primary tumour (Eaton, Tooke et al. 2011). Moreover screening of parents reveals that 32% of the tested parents carried mutation at the same loci. Germline mosaicism has been recently hypothesised, since affected siblings may have negative testing parents (Mobley, McKenney et al. 2010).

Other malignancies have been linked to germline *SMARCB1* mutation, such as Schwannomatosis (Rousseau, Noguchi et al. 2011), chordomas (Mobley, McKenney et al. 2010), Choroid Plexus Carcinoma, Medulloblastoma (Sevenet, Sheridan et al. 1999) and Epithelioid Sarcoma (Modena, Lualdi et al. 2005). The finding of germ-line *SMARCB1* mutation in other embryonal tumours of the central nervous system suggests these malignancies share common pathways of oncogenesis with Malignant Rhabdoid tumours. *SMARCB1* alteration is nowadays used in clinics as a molecular marker to define this spectrum of tumours.

#### *1.1.5.1 Biological and molecular differences between Atypical Teratoid Malignant Rhabdoid Tumours and other embryonal tumours of the CNS*

Embryonal tumours are the most common central nervous system malignancies in infants and young children (typically aged between 0 and 36 months). The World Health Organization (WHO) includes in this classification Medulloblastoma (MB), Supratentorial Primitive Neuroectodermal Tumour (PNET) and Atypical Teratoid Rhabdoid Tumours.

Embryonal tumours present very similar microscopic features, characterised by an undifferentiated small and round stem-like cells. However, pathologically these malignancies have been classified in different subclasses (Figure 1.9): classic Medulloblastoma, desmoplastic Medulloblastoma, Medulloblastoma with extensive nodularity, large-cell variant Medulloblastoma, anaplastic Medulloblastoma, Atypical Teratoid Rhabdoid Tumours and PNET. Despite having similar histological characteristic and location, these malignancies are genetically divergent. Molecular profiling of these malignancies has disclosed their heterogeneity, introducing a more complex and complete understanding of these malignancies.



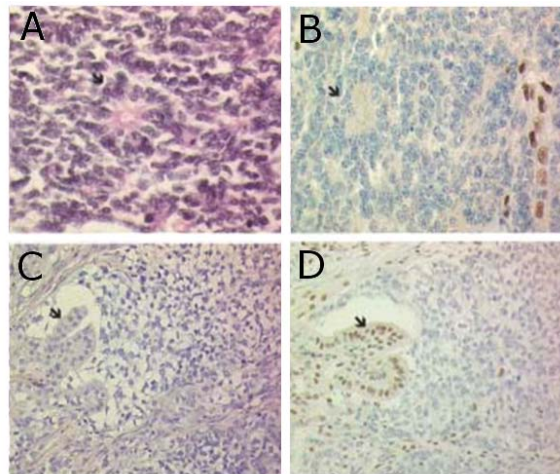
*Figure 1.9 Embryonal tumour microscopy features. A) Classic Medulloblastoma are characterised by a poor differentiated cells. B) Desmoplastic/nodular Medulloblastoma presents reticulum free pale nodules of cells with neuronal or astrocytes differentiation (asterisk). C) Medulloblastoma with extensive nodularity contains cell clustered in nodules with a characteristic elongated shape. (As seen in this panel, both patterns can co-exist in a single tumour. D) Atypical Teratoid/Rhabdoid Tumours shows an undefined appearance; moreover Rhabdoid cells are very sparse. E) CNS primitive neuroectodermal tumours presents anaplastic foci (asterisks) and also undifferentiated cells (Eberhart 2011).*

In particular Medulloblastoma has been divided into four different subgroups: Sonic Hedgehog subgroup (Shh) in which the Sonic Hedgehog signalling pathway drives the tumourigenesis having high levels of expression of *MYCN*; Wnt subgroup characterised by somatic mutations of *CTNNB1* encoding  $\beta$ -catenin; Group 3 and Group 4 have been separated by transcriptional profile, which is for these two groups the gold standard. Moreover, while in Group 3 tumours present frequent amplification of *MYC* Group 4 tumours have low expression of both *MYC* and *MYCN* (Taylor, Northcott et al. 2012).

New studies show that *SMARCB1*-deficiency is not the only biomarker we should use to classify Rhabdoid Tumours and Atypical Teratoid Rhabdoid Tumours. In particular, despite the majority of malignancies presenting homozygous deletion or point mutations associated with loss of heterozygosity (LOH) or double point mutations of the

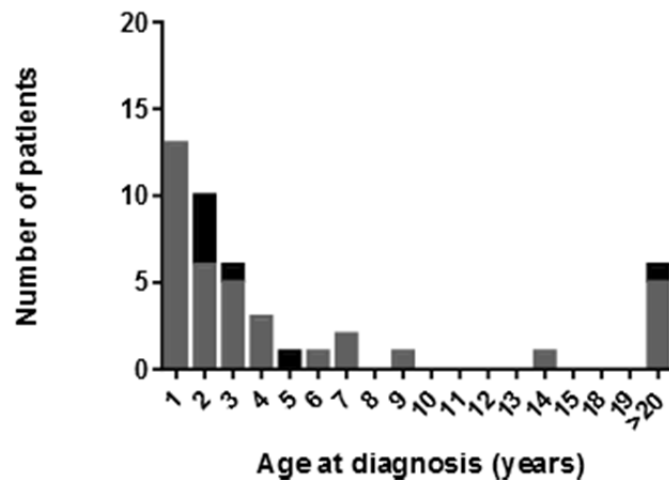


*SMARCB1* gene, rare cases of missense mutation which do affect just the protein function have been found. Moreover, immunohistochemistry alone is not accurate to detect those tumours which have low numbers of Rhabdoid cells. In addition other *SMARCB1*-deficient tumours with a similar location (Choroid Plexus Carcinoma and PNET) of Atypical Teratoid Rhabdoid Tumours have been observed (Figure 1.10)



*Figure 1.10 HES (A-C) and anti-SMARCB1 (B-D) Immunohistochemical staining of SMARCB1 negative tumours non compatible with Rhabdoid morphology. A-B) PNET tumours show typical rosette like structures (arrows) and negativity for anti-SMARCB1; C-D) Choroid plexus carcinomas reveal positivity in epithelial structures (arrows) and negativity for anti-hSNF5/SMARCB1. (Bourdeaut, Freneaux et al. 2007).*

Although *SMARCB1*-negative tumours present morphology incompatible with Atypical Teratoid Rhabdoid Tumours these malignancies have similar poor outcome and are common in children and infants (Figure 1.11) (Bourdeaut, Freneaux et al. 2007).

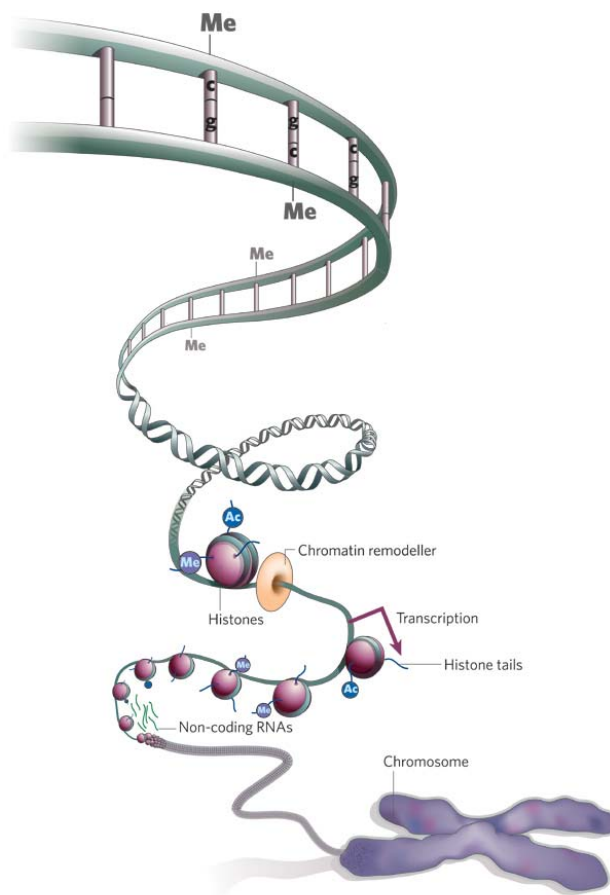


*Figure 1.11 Age at diagnosis of patient with SMARCB1 negative malignancies. The black columns represent tumours with histological features not compatible with Rhabdoid Tumours (Bourdeaut, Freneaux et al. 2007)*

With the application of genetic screening it has become clear that histology alone is inadequate for the diagnosis and classification for CNS tumours. WHO histological classification must be supplemented with genetic analysis that clarifies the molecular difference between these tumours.

## 1.2 Epigenetic changes in cancer

Epigenetic dysregulation plays an essential role in cancer development as well as genetic alterations. Disruption of the epigenetic mechanisms primarily leads to changes in gene expression without alteration in DNA sequence. Alteration in the epigenetic machinery have been attributed to changes in chromatin organisation including DNA methylation, histone modification, nucleosome remodelling and non-coding regulatory RNAs (Figure 1.12) (Jones and Baylin 2002, Herman and Baylin 2003, Portela and Esteller 2010, Sharma, Kelly et al. 2010).



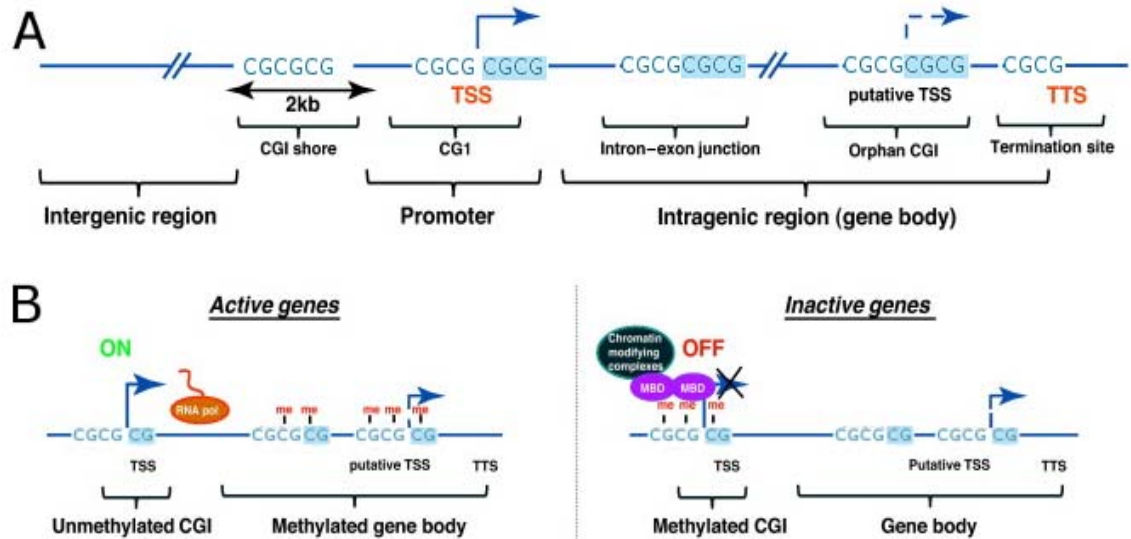
*Figure 1.12 The coding and structural mechanism of chromatin. DNA methylation and covalent histones modifications and chromatin remodeller contribute to nucleosomal remodelling, allowing gene transcription (2008).*

Genetic and epigenetic factors are closely associated, as epigenetic marks and DNA modifications are the direct consequence of interactions between epigenetic modifiers and DNA. Mutation, deletion or the altered expression of one of these epigenetic modifiers can significantly disarrange the epigenetic apparatus and lead to tumour evolution (Esteller 2008). Contrarily, a subset of aggressive malignancies presents a strong genomic stability (like Rhabdoid Tumours) suggesting the involvement of epigenetic alterations in oncogenesis (McKenna, Sansam et al. 2008, McKenna, Tamayo et al. 2012).

### **1.2.1 DNA methylation**

DNA methylation is a chemical modification of the DNA structure that occurs at the cytosine residues in CpG dinucleotide. The cytosine is altered by a family of DNA methyltransferase enzymes (DNMTs) converted to 5-methylcytosine by addition of a methyl group. CpG dinucleotides are usually located in short CpG rich DNA portion called CpG Islands or in regions with highly repetitive sequence (Bogdanovic and Veenstra 2009).

In normal mammalian cells CpG islands are sited in the gene promoters (60%) or in the gene body (Figure 1.13A) (Wang and Leung 2004). Those that occupy promoter regions are usually unmethylated, while those located in the repetitive genomic region are heavily methylated. CpG island promoters show an unmethylated profile especially during development and in undifferentiated tissue (Tsai and Baylin 2011). In contrast hyper-methylation in the repetitive genomic fractions prevents chromosomal instability silencing non coding DNA and Transposable elements (TEs) (Figure 1.13B) (Suzuki and Bird 2008)



*Figure 1.13 Schematic representation of the mammalian DNA methylation mechanism. A) DNA regions characterised by CpG rich portion. CpG island (CGI) promoters, CGI shores (up to 2 kb away from promoters), orphan CGIs and intron-exon. (B) DNA methylation and gene transcription (me=methylation; ON=activation of gene expression; OFF= gene repression). Gene activation is generally promoted by unmethylated CpG islands in the promoter region, associated with methylation in the gene body (left). Methylation of the CpG island on the promoters and consequent demethylation in the body gene is primarily correlated with gene repression (right) (Ndlovu, Denis et al. 2011).*

Abnormality of the methylation machinery is frequent in cancer cells. In particular, gain of methylation at promoter genes has been found to affect very well documented tumour suppressors (Baylin and Herman 2000), causing transcriptional repression and loss of gene function. Genes such as *BRCA1/2*, *PTEN* and *RB* have been found to be hypermethylated and therefore down regulated, leading to a dysregulation of the cell cycle control and DNA repair; this tumourigenic mechanism induces tumourigenesis in a similar way when the same genes have been deleted or mutated (Yoo and Jones 2006). Moreover, changes in promoter methylation of DNA-repair genes such as *MGMT* make

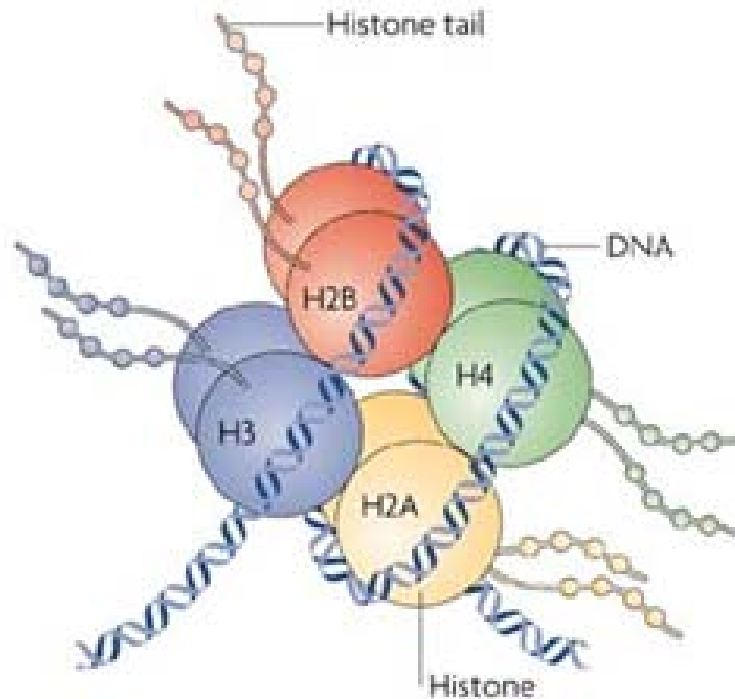
cells more susceptible to mutation of oncogenes such as *TP53* and *KRAS*, as result of genetic instability (Baylin and Jones 2011).

Loss of methylation seems a less frequent event than hypermethylation, however it is strongly associated with early carcinogenesis and tumour progression in chronic myeloid leukaemia, cervical tumours and hepatocellular carcinoma (Ehrlich 2009). Demethylation causes genomic instability and therefore tumourigenesis mainly by promoting chromosomal rearrangement. For example, hypomethylation at the retrotransposon level is responsible for their activation and eventually promotes their displacement to other regions of the genome (Eden, Gaudet et al. 2003, Howard, Eiges et al. 2008). In addition, loss as well as gain of methylation causes activation of genes promoting cell growth (example *R-ras* in gastric samples) and loss of imprinting (*IGF2* in Wilms' tumours) (Sharma, Kelly et al. 2010).

Methylation plays a critical role in carcinogenesis throughout the activation or inhibition of delicate mechanisms that regulate gene expression. Genomic instability and epigenomic modification are strictly connected and the deregulations of these mechanisms lead to tumourigenesis.

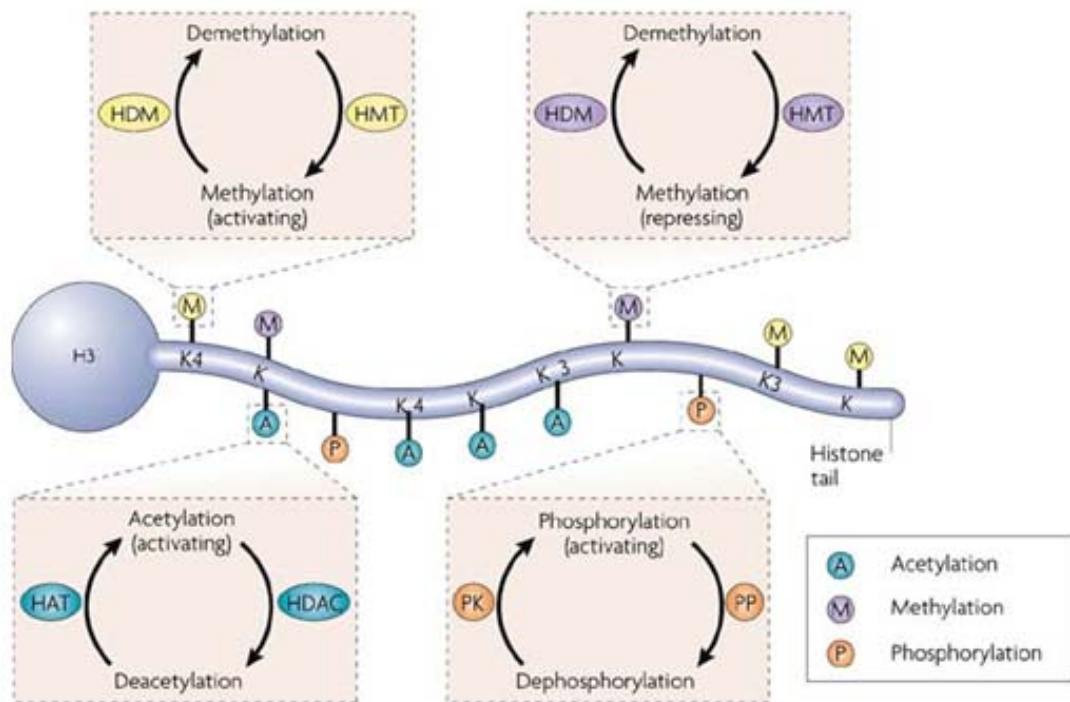
### **1.2.2 Changes in histone modification**

The nucleosome plays an important role in chromatin organization. It is assembled in an octameric structure around which 147 pairs of DNA are tightly wrapped. The protein cores are composed of two copies each of histones H2A, H2B, H3 and H4, which are structurally high conserved (Figure 1.14). Modulation of chromatin dynamics is made possible by specific modification of the histones and in particular by alternating its charge or recruiting other proteins throughout selective recognition.



*Figure 1.14 Structure of a nucleosome: DNA wound around a histone core. (Tsankova, Renthal et al. 2007)*

Histone proteins are characterised by highly basic histone amino (N)-terminal tails, which can be methylated, acetylated, ubiquitylated and phosphorylated following post-transcriptional covalent modification (Figure 1.15). Histone modifications are controlled by enzymes dedicated to introducing or removing acetyl groups (such as histone acetyltransferases -HATs and deacetylase - HDACs) or methyl groups (such as methyltransferases- HMTs and demethylases - HDMs). Histone modification is also driven by Polycomb group proteins (PcGs) and trithorax group proteins (TrxGs). These modifying complexes and the resulting histone conversions play a key role in cellular processes such as transcription, DNA replication and repair by either repressing or activating gene expression.



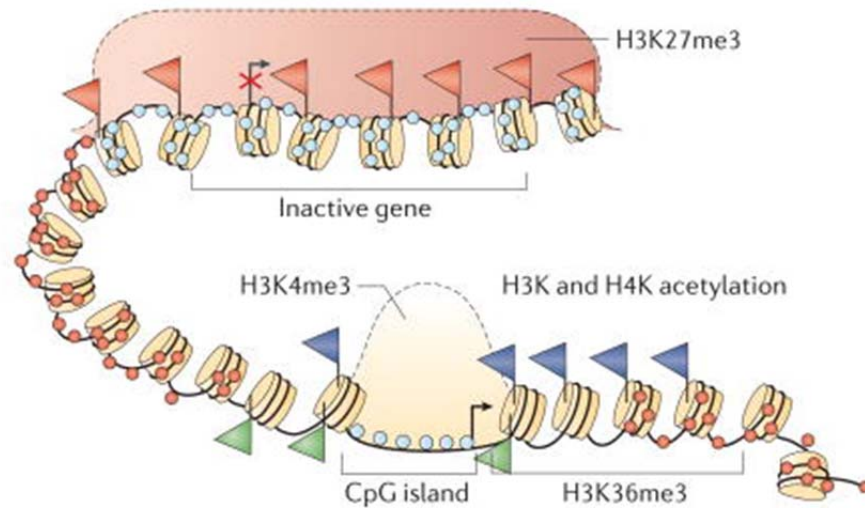
*Figure 1.15 Most common histone modifications, including acetylation, methylation and phosphorylation at several amino acid residues. K= lysine residue; S= serine residue. Acetylation is promoted by histone acetyltransferases (HATs) and reversed by histone deacetylases (HDACs); methylation is maintained by histone methyltransferases (HMTs) and reversed by histone demethylases (HDMs); and phosphorylation is catalysed by protein kinases (PK) and reversed by protein phosphatases (PP) (Tsankova, Renthal et al. 2007).*

The two most common post-translational modifications are acetylation and methylation of histone lysine residues, which are involved in the transcriptional mechanisms of activation of euchromatic state (Jones and Baylin, 2002; Jones and Baylin, 2007; Blackledge and Klose, 2011; Poetsch and Plass, 2011). In particular these histone modifications make DNA accessible to the transcription machinery and other DNA binding proteins, changing the tightly compacted nucleosomes DNA conformation (heterochromatin) into a more relaxed state (euchromatin) (Jones and Baylin, 2002).



In somatic cells the majority of genes present unmethylated CpG islands at their promoters and possibly enhancers; almost 60% of the human genes have nucleosome-depleted regions (NDRs) at the transcriptional start site (TSS). The nucleosomes located nearby NDRs present trimethylation of lysine 4 and 36 of histone 3 (H3K4me3 and H3K36me3) mediated by trithorax group proteins (TrxGs) and acetylation of key lysines of H3 and H4; also they are associated with the histone variant, H2.AZ, which destabilises the nucleosome structures and promotes initiation of transcription (Figure 1.16). Downstream of the TSS, in the repetitive elements and in gene bodies the DNA structure is predominantly methylated and therefore inactive. Promoters are silenced not just through methylation, but also by polycomb group proteins (PcGs). In eukaryotic cells, the complex PRC2 mediates trimethylation of lysine 27 of histone 3 (H3K27me3), which inhibits transcriptional elongation by ubiquitylation and chromatin compaction by recruitment of the complex PRC1 (Muller and Verrijzer 2009); (Stock et al. 2007) (Eskeland et al., 2010).

PcGs and TrxGs interactions are essential in cellular memory system, stem cell renewal, as well as cell homeostasis. Genetic lesions that cause gain of PcGs and loss of TrxG expression has been linked to cancer predisposition by alterations in gene expression level, as well as chromosomal deletion, amplification or translocation (Mills 2010). It is evident that mutations of these enzyme families cause aberrant post-translational histone modifications, thus affecting chromatin dynamics and access to coding, promoter and non-promoter sequences. These enzymes have been often mutated in cancer cells, affecting the delicate transcription machinery and therefore leading to aberrant gene expression. For example, the HATs family are often mutated in colon, uterine, lung tumours and leukaemia (Esteller 2007). Mutation of HDACs has been observed in breast, lung, prostate, gastric and colorectal cancer. Also HMT and HDMC somatic and germline mutation have been reported in a variety of malignancies (Miremadi, Oestergaard et al. 2007).

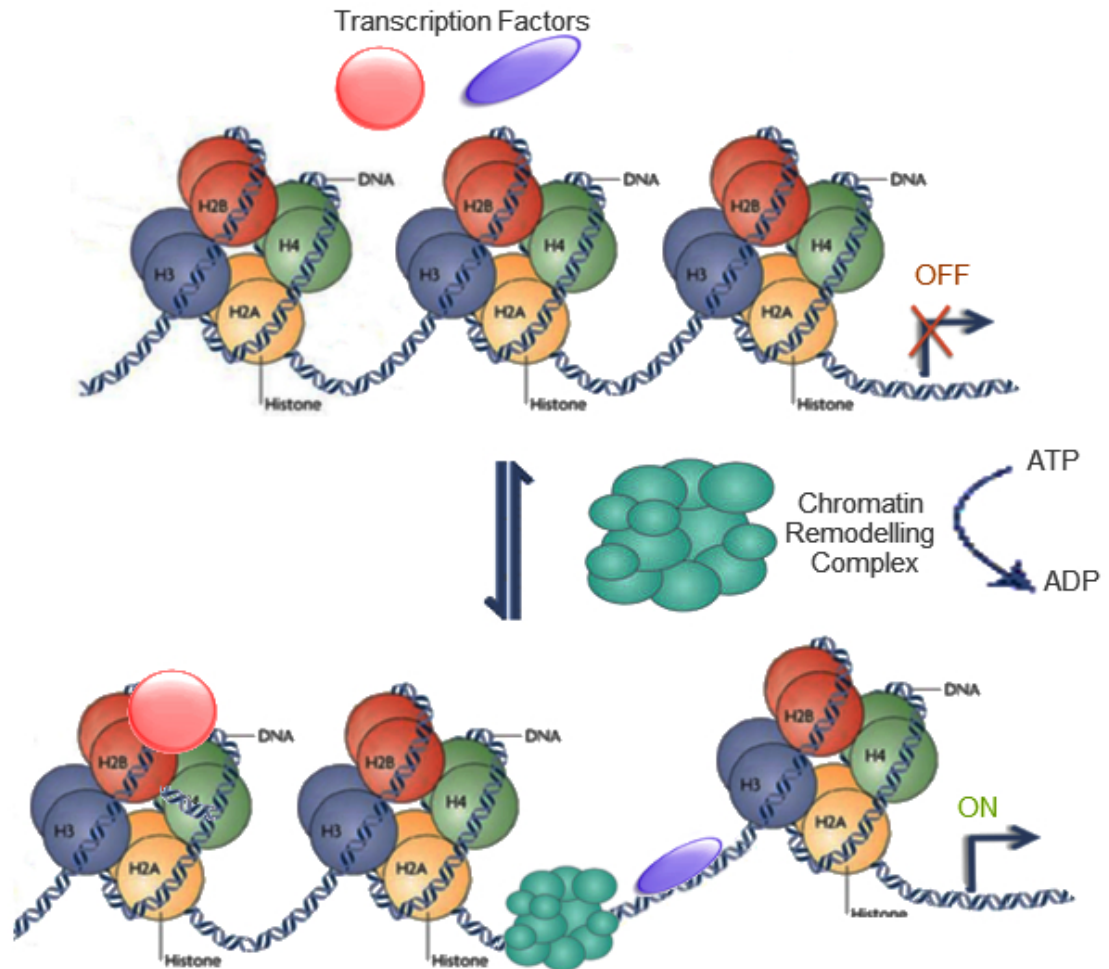


*Figure 1.16 Model of histone modifications in normal human cells. At the top of the figure: silent genes (red X over the transcriptional start site) present promoter region occupied by a Polycomb group complex (represented by a red oval) responsible for repressive histone modification such as trimethylation of lysine 27 on histone 3 (H3K27me3). At the bottom of the figure: genes with fully active transcription state have unmethylated CpG dinucleotides islands within the promoter region (blue pale circles), nucleosomes positioned over the transcriptional start sites and heavily methylated regions upstream of the promoter (red circle); the permissive state is also characterised by active histone H3 lysine 4 and lysine 36 trimethylation (H3K4me3 and H3K36me) and by histones H3 and H4 acetylated at key lysines.*

### 1.2.3 Chromatin remodelling complexes

Nucleosome organisation is a critical process in the epigenetic regulation of gene expression. The nucleosome occupancy is regulated by chromatin remodelling complexes in an ATP-dependent manner. Chromatin remodelling complexes modulate DNA accessibility via two major mechanisms. The first is responsible for transcriptional activation, by the translocation of the nucleosome along the DNA; the second promotes DNA repair, by dimer replacement (Figure 1.17).

Mammalian ATPase chromatin-remodellers have been divided into four main families: the NuRD (nucleosome remodelling and deacetylation)/Mi-2/CHD (chromo-domain, helicase, DNA binding) family, INO80 (inositol requiring 80) family 8 and 11, the SWI/SNF (switching defective/sucrose non-fermenting) family and the ISWI (imitation SWI) family. These complexes present similar ATPase domains but different chromatin interacting domains (chromo, bromo and SANT domain respectively), suggesting a specialised and distinct role in chromatin remodelling (Boyer, Latek et al. 2004, Ho and Crabtree 2010). In particular, NuRD/Mi-2/CHD family are correlated with transcriptional repression and activation of rRNA in the nucleus (Hsu, Lee et al. 2012, Xie, Ling et al. 2012); instead INO80 remodelling complexes activate and repress genes implicated in DNA damage (Van Attikum, Fritsch et al. 2004). The primary function of SWI/SNF and ISW remodelling complexes is to re-organise the nucleosome. However while SWI/SNF complexes promote accessibility for transcription-factor binding and gene activation (Reisman, Glaros et al. 2009, Ronan, Wu et al. 2013), ISW types induce transcriptional repression, activation and elongation (Corona and Tamkun 2004).



*Figure 1.17 Mechanism of transcription activation driven by chromatin remodelling complexes. Remodelling proteins bind to the chromatin using the energy of ATP hydrolysis. The disaggregation of the nucleosome allows the transcription factors to bind and begin transcription.*

Dysregulation of ATPase chromatin remodelling complexes results in tumour development and progression. CHD5 subunit of the NuRD/Mi-2/CHD complexes have been found deleted in Medulloblastoma (White, Thompson et al. 2005), acute myelogenous leukaemia (AML), chronic myelogenous leukaemia (CML), non-Hodgkin's lymphoma, colon cancer, cervix cancer and breast cancer (Bagchi, Papazoglu et al. 2007), indicating a role as tumour suppressor gene. NuRD/Mi-2/CHD

subunit-MTA1 (metastasis –associated gene 1) has been also reported as highly expressed in breast, oesophageal, colorectal and gastric carcinomas (Bowen, Fujita et al. 2004). Several SWI/SNF subunits also are found mutated or deleted in a variety of malignancies. SNF5 subunits undergo a biallelic deletion in Rhabdoid Tumours (Sevenet, Sheridan et al. 1999), renal carcinomas and melanomas (Sevenet, Lellouch-Tubiana et al. 1999, Lin, Wong et al. 2009) and mutations in *ARID1A/BAF250a* have been significantly associated with ER-/PR-/HER2- breast cancers (Zhang, Zhang et al. 2012). In addition, *SMARCA4* ATPase subunit mutations have been identified in lung, Medulloblastoma, Rhabdoid and prostate cancer (Wilson and Roberts 2011).

Reorganisation of the chromatin structure, as well as histone modification regulates a wide range of cellular processes. In particular, ATPase chromatin remodelling complexes regulate transcriptional events, DNA-damage response, DNA replication and cell cycle progression. The mutation of even one subunit of these complexes results in an altered chromatin structure that can initiate tumourigenesis. Despite the evidence implicating ATPase remodelling complexes in tumour development, its role in regulating genome-wide epigenetic regulation still remains to be explained.

### 1.3 SWI/SNF complexes

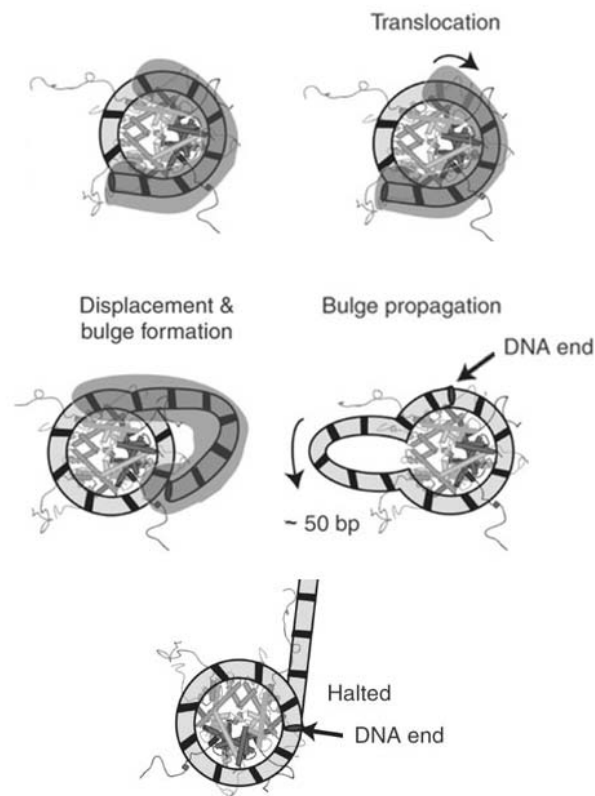
The *SMARCB1* gene encodes SMARCB1 protein which is part of the core of the SWI/SNF chromatin remodelling complex; a modulator of the chromatin structure. The ATP-dependent SWI/SNF complex plays an important role in the remodelling of DNA in order to allow the activation or repression of genes (Utley, Cote et al. 1997)

In mammals the SWI/SNF complexes are evolutionarily conserved and present in diverse isoforms, with different biochemical functions, each of them are formed of 8-10 subunits, including either Brahma (BRM) or Brahma-related gene 1 (SMARCA4), which constitute the ATPase core (Wang, Cote et al. 1996).

The action of the SWI/SNF complex on chromatin remodelling is made possible by the translocation of the nucleosome along the DNA in *cis*. This causes the DNA target sites which were previously occluded by the histone octamer to become exposed, and allows the progression of other nuclear processes (Kingston and Narlikar 1999).

Although the mechanism by which the SWI/SNF complex acts in the remodelling of the chromatin structure is not yet fully explained, two different theories have been proposed. One theory hypothesises that the SWI/SNF complex alters the structure of the DNA using torsional force, by the interaction on the surface of the nucleosome histone core. This torsion results in an alteration of the DNA twist on the surface of the nucleosome allowing the propagation of this twist over the surface of a nucleosome. This mechanism is often referred to as “twist defect diffusion” (Flaus and Owen-Hughes, 2001). A second theory was prompted by the discovery that SWI/SNF complexes are able to modify the chromatin structure by interacting directly with the surface of the histone octamer (Saha, Wittmeyer et al. 2005). The dissociation of the DNA allows the SWI/SNF complex to form a loop of DNA extending out from the nucleosome. It was found that the loops are able to diffuse into the zones next to the core of the histone octamer, shifting the loop in the DNA strand from 5' to 3', like the RNA polymerases (Figure 1.18) (Halliday, Bock et al. 2009). Nowadays this Bulge theory is the most favoured model because large impediments to DNA twisting (hairpin crosslinks), do not prevent nucleosome sliding. The possibility of the SWI/SNF

complex to act on the DNA in the presence of hairpins indicates that that nucleosome sliding does not principally occur following a twisting-diffusion mechanism. In fact the DNA orientation, during the process, does not change: it maintains the same rotational orientation to the histone surface and it appears partially dissociated from the nucleosome (Aoyagi and Hayes 2002).



*Figure 1.18: Scheme of the bulge mechanism. The DNA is wrapped around the core histone protein octamer. The SWI/SNF complex binds to DNA in the weak association between DNA and histone. The SWI/SNF complex forms a loop of DNA of about 52 base pairs. The loop of DNA propagates around the core histone octamer and then released. (Zofall, Persinger et al. 2006)*

### **1.3.1 Role of SWI/SNF complexes in transcriptional regulation**

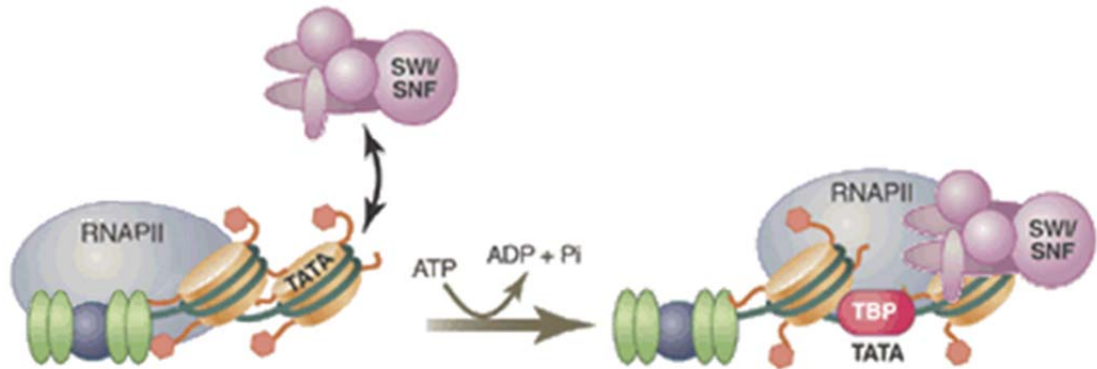
Yeast studies indicate that SWI/SNF complexes regulate approximately 5% of yeast genes (Sudarsanam, Cao et al. 1999). In human cells the SWI/SNF is present at different SWI/SNF-like complexes and this suggests that in humans its gene regulatory function is even more elaborate. Generally SWI/SNF complexes have a role in cell differentiation (Gresh, Bourachot et al. 2005), proliferation (Hah, Kolkman et al. 2010), neural development (Yoo and Crabtree 2009, Weinberg, Flames et al. 2013) and malignant processes.

In eukaryote cells this process is driven by a very complex mechanism, mainly resulting from condensation and relaxation of the chromatin. SWI/SNF complexes play an important role in the dynamic regulation of chromatin structure and consequentially in gene regulation mainly by manipulation of the nucleosome structure (Kadam and Emerson 2003). SWI/SNF complexes drive both repression and activation of lineage and tissue specific genes (Lessard 2007, Kaeser, Aslanian et al. 2008). CD4 and CD8 are both silenced and activated by SMARCA4 and BAF57 during T lymphocyte development (Chi 2002). In embryonic stem cells SMARCA4 not only represses transcription of genes associated with differentiation, but also trigger genes of pluripotency (Ho 2009, Ho and Crabtree 2010).

SWI/SNF complexes are also recruited to individual genes regulating their expression and therefore transcription specificity. Selective recruitment of SWI/SNF to particular promoters requires interactions with specific activator or repressive factors (Ahringer 2000, Sudarsanam and Winston 2000, Kadam and Emerson 2003). Euskirchen et al have tried, using ChIP-seq, to provide a more exhaustive mapping of SWI/SNF binding in Hela cells. In particular, they discovered that the majority of SWI/SNF binding sites are located near protein coding genes ( $\pm 2.5$  kb), which confirms its important role in gene expression. ChIP-seq demonstrated that SMARCB1, SMARCA4, BAF155 and BAF170 frequently bind transcription start sites (TSSs), enhancers, CTCF regions and many regions occupied by RNA Polymerase II and III (Pol II and Pol III) (Euskirchen, Auerbach et al. 2011). This intermediate interaction could facilitate subsequent step



activation, such as recruitment of TATA-binding protein (TBP) at core promoter and transcriptional elongation (Ryan, Jones et al. 1998) (Figure 1.19).



*Figure 1.19 SWI/SNF complexes facilitate activation of gene expression by interaction with TATA-binding proteins. Upstream activators (green and purple) recruit RNA polymerase II (Pol II); subsequently SWI/SNF complex is recruited, disrupting nucleosome structure. This facilitates the binding of TATA-binding proteins (TBP) to the TATA box in promoter region.*

Studies of the contribution of SWI/SNF to gene expression reveal that SMARCA4 and SMARCB1 subunits are necessary and their lack causes a concomitant decrease of RNA levels (Euskirchen, Auerbach et al. 2011, Euskirchen, Auerbach et al. 2012). SMARCA4 regulates the access of STAT proteins to DNA. In particular, in HepG2 cells SMARCA4 binds STAT3 to the *cdkn1A* promoter, causing acetylation of histone H3 (Giraud, Hurlstone et al. 2004). Moreover SMARCA4 is recruited by STAT3 at the IRF1 promoter, indicating a direct involvement of SMARCA4 in the IL-6 pathway (Ni and Bremner 2007). SMARCA4 also plays an important role in pluripotency and self-renewal in stem cells and in particular is involved in blastocyst development occupying pluripotency genes, lineage-specific genes, and epigenetic modifier genes (Kidder, Palmer et al. 2009). Interestingly SMARCA4 is required at the promoter of genes that are essential for pluripotency in the stem cell, such as *Oct4*, *Sox2* and *Sall4*, suggesting

that SMARCA4 is directly involved during preimplantation development but only after the establishment of pluripotency (Kidder, Palmer et al. 2009). SMARCB1 acts as transcriptional cofactor and is essential for recruitment of SMARCA4 at the promoter region of genes such as *CDKN2A*, *CDKN1A* and *Cdk* (Betz, Strobeck et al. 2002, Chai, Charboneau et al. 2005). Although SMARCB1 is mainly involved in gene activation, it also contributes to the mechanism of repression; in particular SMARCB1 represses *Cyclin D1* and *Aurora A*, in a histone deacetylase dependent manner by recruitment of HDAC1 (Zhang, Davies et al. 2002, Lee, Cimica et al. 2011).

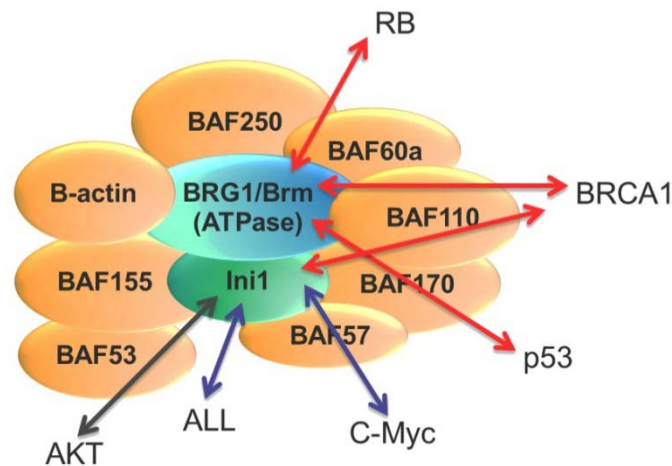
These results evidence the recruitment of SWI/SNF complex promoted by *SMARCB1* to induce gene expression. However *SMARCB1*'s role in gene expression remains still unclear; in particular how *SMARCB1* loss affects the SWI/SNF complex and therefore gene expression.

#### *1.3.1.1 The critical role of the SWI/SNF complex in cancer development*

The SWI/SNF complex regulates a variety of biological activities, such as DNA methylation, repair, recombination and replication (Narlikar, Fan et al. 2002). Loss of function can easily contribute to malignant formation (Bultman, Gebuhr et al. 2000, Glaros, Cirrincione et al. 2007). Moreover, it has been recently discovered that the SWI/SNF complex interacts with proteins produced by oncogenes (Figure 1.20), such as c-Jun (Ito, Yamauchi et al. 2001), MYC (Cheng, Davies et al. 1999) and MLL (Zuqin *et al*, 2003) and also with tumour suppressor proteins such as *pRb* (Dunaief, Strober et al. 1994, Trouche, Le Chalony et al. 1997, Strobeck, Knudsen et al. 2000) *TP53* (Lee, Kim et al. 2002) (Jia, Osada et al. 1997) *BRCA1* (Jia, Osada et al. 1997, Biegel 1999) and *AKT* (Forest, David et al. 2012).

The SWI/SNF complex also influences cellular differentiation and proliferation, controlling the cell-cycle using two different pathways. Its ATPase subunits BRM and in particular SMARCA4 are recruited by the Rb protein, inducing growth arrest at G1-S (Trouche, Le Chalony et al. 1997, Strobeck, Knudsen et al. 2000). Dominant negative mutant SMARCA4 cell lines show resistance to growth arrest, caused by a disruption of the RB pathway: SMARCA4 is necessary to mediate RB signalling to Cyclin A/Cdk2,

resulting in G1 arrest (Strobeck, Knudsen et al. 2000) Moreover, Rb interacts with histone deacetylase (HDAC) and the hSWI/SNF, resulting in an inhibition of transcription of genes coding for cyclins E and A and in the arrest at the G1 phase of the cell cycle (Zhang, Davies et al. 2002).



*Figure 1.20: Structure of the SWI/SNF complex figure shows direct interaction between SMARCB1 and SMARCA4/BRM with tumour suppressor proteins (red), oncogenes (blue) and AKT.*

Experiments by Lee *et al* suggest that the SWI/SNF complex is necessary for proper induction of the p53 pathway (Lee, Kim et al. 2002). For instance, studies of dominant negative mutations of SMARCA4 and *SMARCB1* null cell lines show a decrease in the expression of *p53*, preceded by a down regulation of *p21*. This complex plays an important function in DNA repair, too. Also it was discovered that, following damage to DNA, the SWI/SNF complex is recruited in order to facilitate the nucleotide excision repair and the double-stranded break repair (Halliday, Bock et al. 2009).

### **1.3.2 SWI/SNF subunits and tumourigenesis**

#### *1.3.2.1 SWI/SNF subunits are frequently mutated in a variety of tumours*

The increased study of these complexes has disclosed an extreme polymorphic structure with specialised function in specific tissues (Wang, Cote et al. 1996). The discovery of *SMARCB1* biallelic mutation in Rhabdoid Tumours (Versteeg, Sevenet et al. 1998, Biegel 1999) has identified that SWI/SNF complexes acting as tumour suppressors. Kadoch and colleges showed that in human tumours SWI/SNF subunits are more frequently mutated than all other chromatin remodelling complexes, suggesting an important role in tumourigenesis (Kadoch, Hargreaves et al. 2013).

Recent genome wide sequencing showed that other SWI/SNF subunits are implicated in tumourigenesis and now it is recognised that 20% of all human tumours contain mutation in at least one of the subunits (Figure 1.21) (Kadoch, Hargreaves et al. 2013). Interestingly, *SMARCB1* is not one of the most mutated subunit. Nevertheless a wide variety of solid tumours show mutation of SWI/SNF subunit genes *ARID1A* and *SMARCA4*, with a high incidence in aggressive malignancies such as clear cell ovarian cancer, gastric cancer, Medulloblastoma and melanoma (Biegel, Busse et al. 2014).

Few malignancies show a complete loss of heterozygosity (LOH) of one of the SWI/SNF gene subunits. In fact 50% of the cases present multiple mutations at the SWI/SNF subunits, such as missense mutation or gene fusion in a variety of different tumours (carcinoma renal, clear-cell ovarian, colorectal carcinoma, gastric cancer melanoma lung cancer, Glioma, breast cancer) (Biegel, Busse et al. 2014). The presence of other mutations in more than one SWI/SNF subunit might disrupt the complex functionality, as results resulting of compound heterozygosity (Kadoch, Hargreaves et al. 2013).

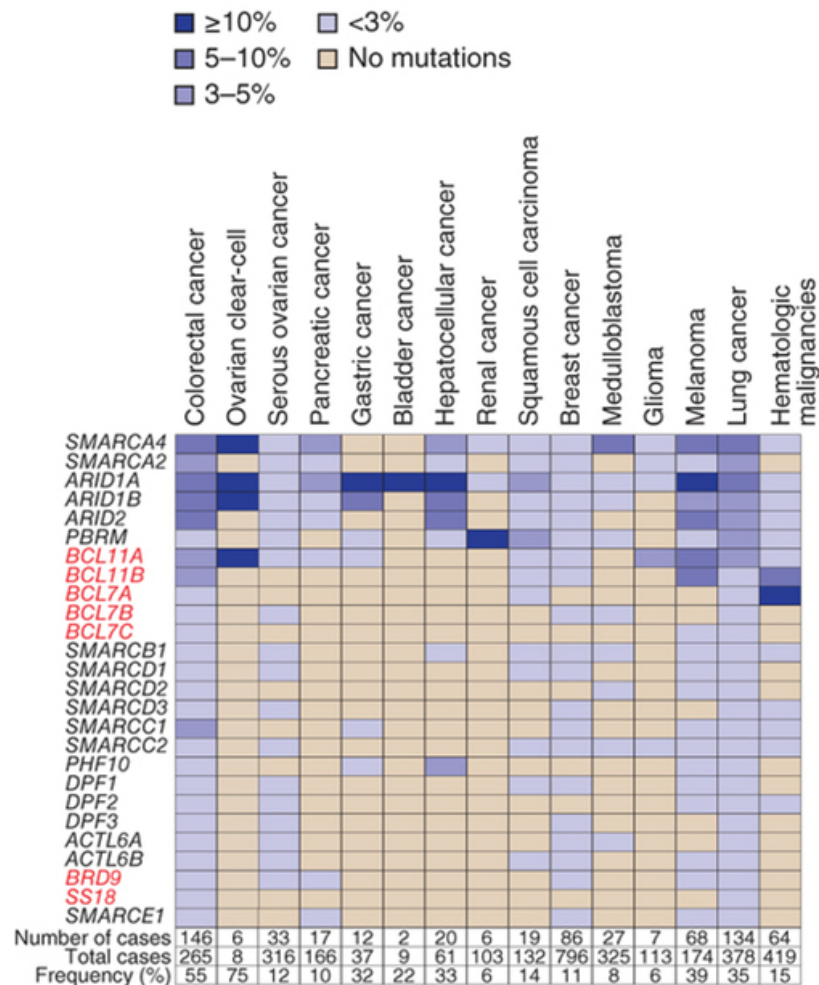


Figure 1.21 Multiple mutations of the mammalian SWI/SNF subunits can occur in human cancers. Frequency of patients harbouring a mutation is presented according to a graded colour scale: in dark blue is indicated frequency of patients harbouring a mutation equal or greater than 10%, in pale yellow no mutation detected. A variety of aggressive malignancies have mutation of ARID1A and SMARCA4 subunits genes.

SWI/SNF complexes have a highly specialised activity to unwrap the chromatin, allowing gene regulation. The selectivity of the SWI/SNF complex is due to their extremely complicated structures, whereby every subunit has a synchronous but specialised role. Co-occurring mutations in multiple subunits or a complete deletion of

just one component of these complexes disrupts a delicate balance and finally leads to tumourigenesis.

#### 1.3.2.2 *The ATPase subunits of SWI/SNF complex: role in tumourigenesis*

SWI/SNF is characterised by the presence of two alternative ATPase subunits, which are able to hydrolyse ATP to ADP namely SMARCA4 and BRM. These subunits do not co-exist in the same complexes and play differential roles in cellular proliferation and differentiation (Phelan *et al*, 2009). Although mammalian SMARCA4 and BRM differ only in the first 60 amino-acids sharing an approximate identity of 75% at the protein level, they have different biological functions in animal models (Muchardt and Yaniv 1999). In particular, in mice SMARCA4 knock-out leads to embryonic death at the peri-implantation stage (Bultman, Gebuhr *et al*. 2000). Other experiments show that tumour cells with a deletion of *BRM* are able to activate alternative mechanisms for controlling the cell cycle (Medina and Sanchez-Cespedes 2008). However there is the possibility that SMARCA4 and BRM may compensate for each other in some cases (Coisy-Quivy, Disson *et al*. 2006). *BRM* null mice are characterised by a level of expression of *SMARCA4* threefold higher than the wild type *BRM* animal (Reyes, Barra *et al*. 1998).

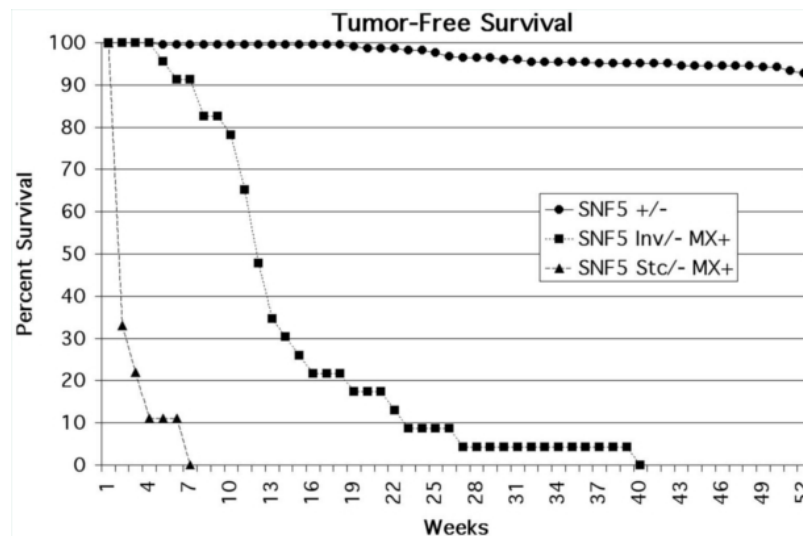
It was demonstrated that a loss of *BRM* causes a modification to the cell cycle that not even over-expression of *SMARCA4* can compensate for. In particular, in 3T3 fibroblasts lacking *BRM* a delay of S phase, an increase in the duration of the M phase and frequent faulty chromosome segregation events in anaphase were observed (Coisy-Quivy, Disson *et al*. 2006). Moreover studies in *BRM* null mice reveal that they are predisposed to develop tumours; knock-out *BRM* mice present about 10 times more lung tumours larger than that of the wild type models. However, the absence of BRM does not modify the penetrance of tumours following *SMARCA4* alteration, but may change the types of tumour that occur. These data suggest that *BRM* and *SMARCA4* could be involved in different stages of tumourigenesis, specifically tumour Initiation and tumour progression respectively (Glaros, Cirrincione *et al*. 2007).

Analysis in a number of lung cancer and other cell lines reveal that expression of BRM and SMARCA4 is lost in about 30-40% of cases however the loss of both subunits has a low frequency (Reisman, Glaros et al. 2009). Decristofaro *et al*, also report the loss of both SMARCA4 and BRM protein expression in 10% of established tumour cell lines (Adenocarcinoma of Kidney Adrenal Cortex, Pancreas, Ovary, Prostate and colon; Breast Tumour; Rhabdoid and Medulloblastoma) (Decristofaro, Betz et al. 2001). In addition, a loss of expression is found in about 15-20% of primary non-small-cell lung cancers. In 10-20% of all bladder, colon, breast, melanoma, oesophageal, head/neck, pancreas and ovarian cancer immunohistochemical staining has demonstrated a loss of expression of both *SMARCA4* and *BRM* (Reisman, Glaros et al. 2009).

#### 1.3.2.3 *The role of SMARCB1 in Malignant Rhabdoid Tumours*

The role of *SMARCB1* (also known as *BAF47* and *hSNF5*) in Malignant Rhabdoid tumours tumorigenesis was recently discovered (Versteeg *et al*, 1998). Most Atypical Teratoid Rhabdoid Tumours have biallelically inactivated their *SMARCB1* genes either by deletion of part of or the entire gene or by point mutation. Moreover the identification of *SMARCB1* germline mutation indicates this gene as a classical tumour suppressor gene (Sevenet, Sheridan et al. 1999, Biegel 2006).

The tumour-suppressing role of *SMARCB1* was studied in mice. Homozygous deletion of *SMARCB1* in mice causes death in the peri-implantation stage of gestation (Guidi, Sands et al. 2001). Instead *SMARCB1*<sup>+/-</sup> mice develop normally, however 25% of them develop Rhabdoid Tumours and undifferentiated sarcomas, with aggressive and metastatic characteristics primarily on the head and neck (Klochendler-Yeivin, Fiette et al. 2000). In addition, 100% of the conditional *SMARCB1* knock-out mouse models developed malignancies (mature CD8+T cell lymphoma or Rhabdoid Tumours ) with a median onset of only 11 weeks (Figure 1.22) (Roberts, Leroux et al. 2002).



*Figure 1.22 Survival of SMARCB1 target mice. SMARCB1<sup>+/-</sup> mice show predisposition to develop Rhabdoid Tumours. Mx<sup>+</sup> SMARCB1<sup>Cond/-</sup> mice rapidly develop bone marrow and organ failure following injection with Polyinosinic-polycytidylic acid (Poly(I)/Poly(C)) and eventually die. Overall 90% of the mice die within 3 weeks. Mx<sup>+</sup> SMARCB1<sup>Inv/-</sup> mice show high tumour penetrance (100%) with average life time of 11 weeks after injection with Poly(I)/Poly(C) (Roberts, Leroux et al. 2002).*

The accumulation of phosphorylated p53 and increased aberrant mitotic features following *SMARCB1* inactivation in murine embryonic fibroblasts indicate that *SMARCB1* has an important role to play in the DNA damage response. 100% of *SMARCB1*<sup>+/-</sup> *p53*<sup>-/-</sup> mice develop tumours within 19 weeks; instead just 28% of the *SMARCB1* <sup>+/-</sup> mice develop tumours with a mean latency of 42 weeks (Klochendler-Yeivin, Picarsky et al. 2006). This indicates a strong genetic interaction between these two tumour suppressor genes.

Vesteege (2002) demonstrated that reexpression of wild type *SMARCB1* in Malignant Rhabdoid tumours cells leads to the arrest of the cell cycle in G1, blocking entry into the S phase. The result of cell cycle arrest is the appearance of flattened cells and the activation of senescence-associated protein, down-regulation of *E2F* (activator of the



transcription of cyclins) and of *Cyclin D1*, activation of Rb and up regulation of *CDKN1A* and *CDKN2A* (Versteeg, Medjkane et al. 2002, Chai, Charboneau et al. 2005). Interestingly, studies in cell lines lacking pRB function have showed that it is not possible for *SMARCB1* to arrest the cell cycle in G1 phase. This underlines the complementary role of *SMARCB1* and *pRb* in the cell cycle to induce a G1 arrest in Malignant Rhabdoid tumours cells (Versteeg, Medjkane et al. 2002). Moreover studies in *SMARCB1* deficient Malignant Rhabdoid tumours cell lines show that the reexpression of *Rb* or *CDKN2A* induces a cell cycle arrest, indicating that *SMARCB1* regulates *pRb* and *E2F* target genes (Versteeg, Sevenet et al. 1998, Versteeg, Medjkane et al. 2002)

*SMARCB1* loss directly disrupts SWI/SNF activation of promoter regions of various genes involved in different pathways. ChIP assays performed by Zhang *et al* showed that in Rhabdoid MON cells *SMARCB1* is directly recruited to the *Cyclin D1* promoter, and the absence of *SMARCB1* causes an increase in binding to the *Cyclin D1* promoter DNA as observed by CHIP. Interestingly chip assays have indicated that *SMARCB1* does not interact directly with pRb but its activity is linked to the recruiting of *CDKN2A* (Oruetebarria, Venturini et al. 2004): a *SMARCB1* binding site was localised to the p16INK4a promoter, suggesting that its presence is correlated with transcriptional activation. Expression analysis has also demonstrated that in Rhabdoid cell lines the transcription of *CDKN1A* is correlated with the enforced re-expression of *SMARCB1* (Zhang, Davies et al. 2002, Kang, Cui et al. 2004). Interestingly, *SMARCB1* re-expression not only increases the binding to *CDKN2A* and *CDKN1A* promoters, but also evicts the polycomb complex silencing proteins such as BMI-1, followed by H3K4 methylation and H3k27 demethylation (Kia, Gorski et al. 2008, Kuwahara, Charboneau et al. 2010, Knutson, Warholic et al. 2013). Recently, chip analysis revealed that in Rhabdoid cells both *in vitro* and *in vivo*, the loss of *SMARCB1* causes an aberrant activity of Aurora A: the re-expression of *SMARCB1* in Rhabdoid tumour cells causes *SMARCB1* to bind to the Aurora A promoter sequence, resulting in a down-regulation of Aurora A and subsequent changes in cell cycle arrest (Lee, Cimica et al. 2011).

Despite a number of studies re-expressing *SMARCB1* in Rhabdoid cells, the mechanisms of tumorigenesis are still poorly understood and results suggest that the

role played by this gene is highly complex. In fact, re-expression of *SMARCB1* does not lead to the same effects on *CDKN2A*, *CDKN2A* and Cyclin D1 in all Malignant Rhabdoid tumours cell lines (Versteeg, Medjkane et al. 2002, Kuwahara, Charboneau et al. 2010, Kuwahara, Wei et al. 2013). Moreover, the lack of *SMARCB1* subunits in SWI/SNF doesn't compromise the chromatin remodelling activity (Phelan, Sif et al. 1999), indicating that its abnormal action in Malignant Rhabdoid tumours could be caused by an aberrant action of *SMARCA4*. It was also demonstrated that the simultaneous deletion of *SMARCA4* and *SMARCB1* in mice prevents the development of tumour (Wang, Sansam et al. 2009), instead in Malignant Rhabdoid Tumour cells knocking out *SMARCA4* causes a dramatic growth arrest (Betz, Strobeck et al. 2002, Versteeg, Medjkane et al. 2002, Zhang, Davies et al. 2002, Doan, Veal et al. 2004).

#### *1.3.2.4 SMARCA4 subunit in SMARCB1 deficient cells: loss of a delicate balance*

This evidence suggests that in Malignant Rhabdoid tumours tumourigenesis may be caused not simply by the loss of *SMARCB1* gene, but also by an aberrant activity of the ATPase subunits of the SWI/SNF complex as result of the *SMARCB1* deletion. In fact the loss of *SMARCB1* may lead to unregulated activity of the SWI/SNF complex that may modify the expression of gene normally involved in the senescence and in the control of the cell cycle.

Even without *SMARCB1* subunits the SWI/SNF complexes are still capable of remodelling chromatin. This is due to the fact that the SWI/SNF complex does not require an intact structure: even in the absence of additional subunits *SMARCA4*, *BRM*, *BAF155* and *BAF170* are characterised by an activity comparable with that of the whole SWI/SNF complex (Phelan *et al*, 2009).

Deficiency of *SMARCB1* leads to a mis-regulation of the SWI/SNF complex, even if the genome-wide consequence of this is still largely unknown. It has been demonstrated that the *SMARCB1* loss causes a deregulation of the epigenetic transcriptional regulatory activities of the SWI/SNF complex and contributes to cancer Initiation. In

Malignant Rhabdoid cells tumour development may be caused by an aberrant activity of ATPase of the SWI/SNF complex, due to an incorrect expression of the SWI/SNF-dependent genes (Hasselblatt, Gesk et al. 2011). Moreover, new studies in *SMARCB1*-deficient Rhabdoid cell lines have indicated that knock-down of SMARCA4 causes cell cycle arrest at G0-G1, ultimately resulting in cell death (Wang, Sansam et al. 2009). Similar results have been detected upon the reintroduction of *SMARCB1* into Malignant Rhabdoid Tumour cells (Betz, Strobeck et al. 2002, Versteeg, Medjkane et al. 2002, Zhang, Davies et al. 2002). Interestingly, *in vivo* a suppression of SMARCA4 in *SMARCB1* deficient mice causes a decrease of tumour development underlining the participation of this subunit in carcinogenesis (Wang, Sansam et al. 2009). This suggests that in *SMARCB1*-deficient cell lines an inhibition of the ATPase subunit SMARCA4 may be sufficient to inhibit tumour progression.

## **1.4 Next generation sequencing and array-based analysis techniques: new techniques in tumour characterisation**

### **1.4.1 Next generation sequencing**

The introduction of the Maxam-Gilbert method in the 1970s opened the field of genomic analysis by DNA sequencing. In 1977 Frederick Sanger developed a more accurate method based on the selective incorporation of chain-terminating deoxynucleotides by DNA polymerase during *in vitro* DNA replication (Sanger, Nicklen et al. 1977). The Sanger sequencing method is still used in most medical and research laboratories. In 1990 the Human Genome Project (HGP) was launched with the purpose of sequencing the entire length of the human genome and identifying and mapping all of the genes within it. The study was released in 2003, resulting in high-throughput sequencing of 99% of the human euchromatic portion (Schmutz, Wheeler et al. 2004).

HGP developed the first generation sequencing method, based on Sanger method. Although at the time Sanger sequencing was considered gold standard for nucleotide sequencing, this approach showed limitations in accuracy, speed and cost. Over the past 10 years DNA sequencing methods have improved, resulting in a much cheaper and higher throughput alternative to the traditional Sanger sequencing.

The fast development and wide applications of next-generation sequencing (NGS) technologies provide more complete and accurate genomic information. Thus improving not just knowledge in translation genomics, but also a better understanding of cancer and other complex diseases, leading to improvement of diagnostics, prognostics, and therapies (Grada and Weinbrecht 2013). In fact one of the most promising potential applications of genome-wide studies is identification of gene signatures that can improve not just the clinical parameters but also the understanding of pathways involved. In fact NGS have high sensitivity and also give much longer sequences reads than other methods.

#### **1.4.2 Transcriptome profiling: RNA sequencing**

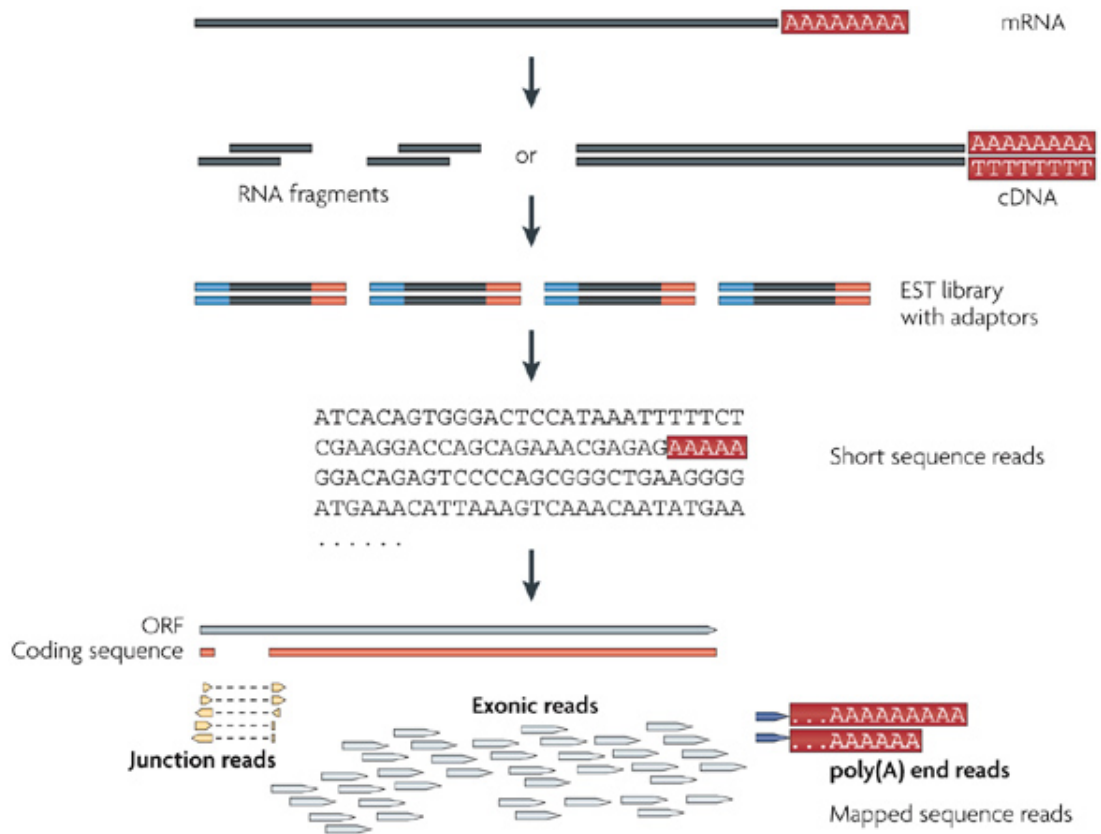
Development of NGS allows more comprehensive studies of all components of the genome, including transcriptome, as a complete set of transcripts in a cell. The transcriptome is the result of genome expression and consists of all the RNA molecules (including mRNAs, non-coding RNAs and small RNAs) produced in a cell or collection of cells. Study of the transcriptome is essential to understand the contribution of alteration expression of genetic in development and diseases such as cancer (Brown 2002). In the past few years application of NGS methods, such as RNA-sequencing (RNA-seq) in transcriptome analysis has allowed detection of differentially expressed genes in disease and even detection of splicing variants gene polymorphism (de Magalhaes, Finch et al. 2010, Zhao, Fung-Leung et al. 2014).

RNA-seq technology has considerable advantages when performing transcriptome profiling, detecting novel transcripts, and allele-specific expression and splice junctions. In contrast, to older techniques RNA sequencing is a sequence-based and tag-based approach to determine directly the cDNA sequence, improving sensitivity to low

abundance transcripts (Brown 2002). In particular, species- or transcript-specific probes are not used in RNA-Seq providing far more precise gene expression levels. Moreover, this approach allows detection of novel transcripts and other changes, such as single nucleotide variants, insertions and deletions, gene fusions and novel transcripts (Grada and Weinbrecht 2013).

The most widely used NGS platforms for RNA-seq are SOLiD, Ion Torrent (Life Technologies) and HiSeq (Illumina). The platforms differ in their library preparation while sample input requirements and the number of cDNA fragments sequenced per run are similar. HiSeq and Ion torrent sequence using a “sequencing by synthesis” (SBS) approach, which allows base-by-base sequencing; instead SOLiD adopt a “sequencing by ligation”, using a DNA ligase (Liu, Li et al. 2012). NGS platforms share a similar workflow that involves template preparation, library preparation and amplification sequencing and data analysis (Figure 1.23) (Wang, Gerstein et al. 2009).

High-throughput sequencing of RNA opens the possibility of analysing millions of RNA transcripts and isoforms within a single assay. For this reason RNA-seq is becoming more appealing to clinical application as diagnostic and prognosis tests (Meldrum, Doyle et al. 2011, Van Keuren-Jensen, Keats et al. 2014). At present RNA-seq is extensively used in research, especially in understanding the association between variations in the transcriptome and tumourigenesis. In-depth quantitative studies in a variety of cancers identify novel mutations, disclosing the benefit of this technique in disease classification by exome sequencing. Massive parallel sequencing provides in-depth analysis of the genome, allowing also characterisation of promoter usage, rare splicing events, epitranscriptome modifications, allele-specific expression or nonsense-mediated decay (Wang, Gerstein et al. 2009, Meldrum, Doyle et al. 2011).



*Figure 1.23 RNA workflow. The samples are prepared by direct fragmentation of cDNA or RNA in order to obtain small fragments (200-500bp). Each cDNA fragments are amplified by PCR and adapters (blue) are added to create the RNA-seq library for directional sequencing. Short sequence reads are aligned with a genome of reference or a de novo assembled transcript; the reads can generally be mapped to four categories: exonic, partially overlaps exon, intronic or between genes.*

### 1.4.3 DNA-protein interaction: ChIP sequencing

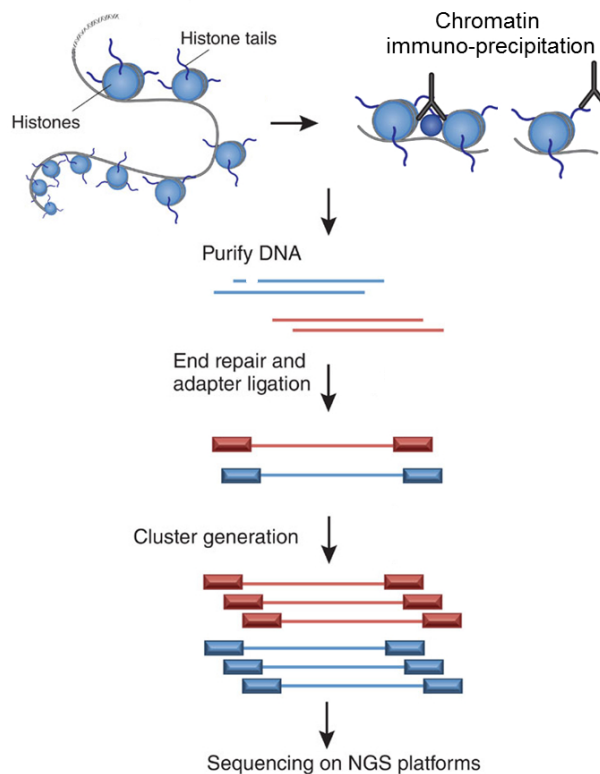
For years understanding of transcriptional regulation has been based on characterisations of genetic modulation of specific genes or gene families. However, the observation of changes in gene expression without alteration in DNA sequence has revealed the importance of epigenetic mechanisms in gene regulation (Baylin and Jones 2011). In particular it has become clearer that analysis of the chromatin state and the DNA-protein interaction is essential to fully understand transcriptional modulation (Park 2009).

Chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) nowadays is the standard technique to investigate and study transcription factor binding, histone modifications (Northrup and Zhao 2011). ChIP-Seq is a modification of ChIP-Chip (hybridization microarray based technique), consisting of a combination of chromatin immunoprecipitation and next-generation sequencing of extracted DNA fragments (Johnson, Mortazavi et al. 2007, Robertson, Hirst et al. 2007). ChIP-seq allows us to assay *in vivo* protein-DNA binding on a genome scale. In fact, the direct sequencing of the DNA fragment results in a higher resolution, the possibility to better filter out artefacts and higher coverage than the earlier array methods (Figure 1.24) (Johnson, Mortazavi et al. 2007, Park 2009, Ashoor, Hérault et al. 2013).

ChIP-seq is becoming popular technique, especially after ENCODE project has been launched in 2003. This project aims to characterise not only specific DNA-protein and RNA interaction sequences, but also those regulatory elements on the DNA responsible for gene expression modulation. Moreover, the rapid development of the next generation sequencing platform has contributed to make ChIP-seq a more robust technique, increasing precision and replicability (Park 2009).

Recently studies show ChIP-seq approaches provide additional information, complementary to those collected by mRNA-level measurements. Combination of data between NGS techniques allows a better understanding of mechanisms involved in gene expression regulation (Gilchrist, Fargo et al. 2009, Gupta, Wikramasinghe et al. 2010, Tummala, Mali et al. 2010, Mokry, Hatzis et al. 2012). In the past few years ChIP-seq

has been used widely to characterise genome-wide transcription factor-binding site in a variety of tumours. For instance, this technique was applied to breast pathology tissue allowing a better definition of the role of epigenetic alterations in breast cancer and characterisation of useful biomarkers (Fanelli, Amatori et al. 2010, Ross-Innes, Stark et al. 2012).



*Figure 1.24 ChIP-seq work flow. Proteins are cross-linked to their bound DNA by formaldehyde treatment and chromatin is sheared and immunoprecipitated with antibody-bound magnetic beads. The immunoprecipitated DNA is used to create a next-generation sequencing library. Specifically a library is created by addition of single adenosine residues, adaptor ligation and PCR with primers compatible with the sequencing platform. Following cluster generation, sequencing is performed on next-generation sequencing platforms (Kidder, Hu et al. 2011).*



#### **1.4.4 Array-based analysis techniques: Methylation array**

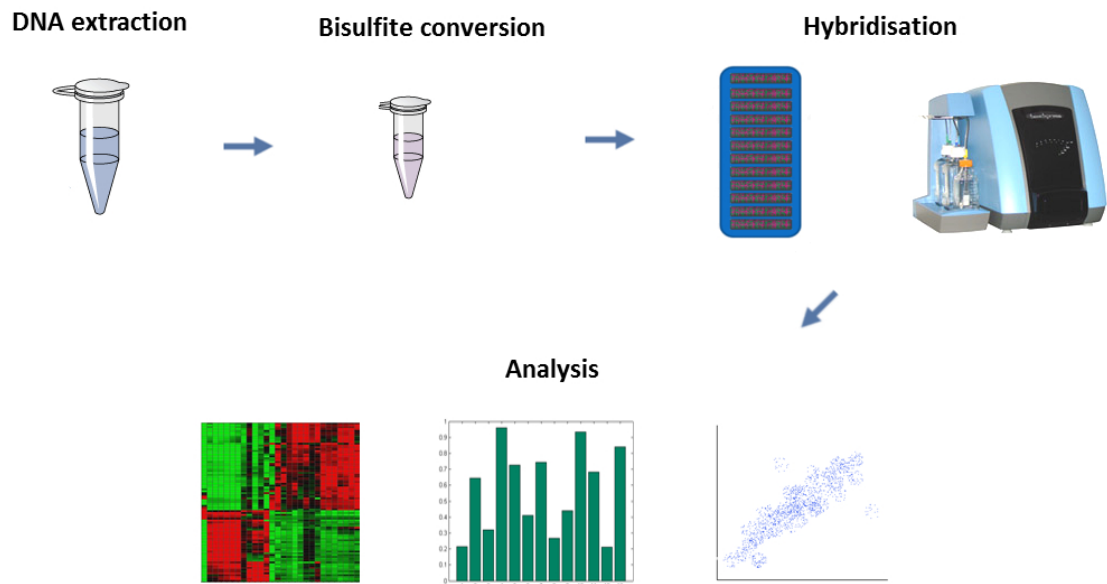
In the past decade studies on DNA methylation and its implication in human diseases have grown dramatically. Alteration of the epigenetic machinery in humans is often associated with aberrant methylation, especially resulting in the development and progression of cancer. The implication of DNA methylation in regulating gene expression has led to a progressive development of commercial epigenetic profiling platforms.

Currently whole genome arrays are available, providing genome-wide epigenetic profiling analysis. The most widely used arrays rely on bisulfite conversion to measure methylation of CpG dinucleotides. Other options for measuring methylation include, whole-genome bisulfite sequencing, restriction enzyme-enriched sequencing, affinity - enrichment-based sequencing, and locus-specific methylation analysis i.e. through bisulfite Sanger sequencing. All of these approaches are based on sodium bisulfite conversion, enzyme restriction digestion or affinity enrichment.

CpG-specific arrays constitute the most widely used technique. In particular, the new generation array Infinium Human Methylation 450K BeadChip (Figure 1.25) allows coverage of 99% of all NCBI Reference Sequence (RefSeq) and assesses the methylation status of more than 480,000 CpG overall. Comparison with previous technologies shows that 450K methylation array has higher accuracy and reproducibility (Bibikova, Barnes et al. 2011, Sandoval, Heyn et al. 2011). This array is currently used in epigenome-wide association studies (EWAS) and in the International Cancer Genome Consortium (ICGC) project.

Currently most of the oncology clinical tests are based on cytogenetic analysis to detect genetic aberrations (Cottrell, Al-Kateb et al. 2014). The limited information provided by these diagnostic tests has opened a discussion to introduce methylation array-based analysis in clinics. In fact, methylation status analysis can disclose candidate biomarkers (Lange, Campan et al. 2012, Naumov, Generozov et al. 2013, de Araujo, Marchi et al. 2014, Karlsson, Jonsson et al. 2014, Klajic, Busato et al. 2014, Li, Li et al. 2014, Yu, McIntosh et al. 2014, Zhang, Xing et al. 2014) which can improve the understanding of

cancer heterogeneity and mechanisms of drug resistance. Moreover, the development of new software for analysis of Infinium Human Methylation 450 BeadChip has improved its use, especially for large-scale clinical studies (Wu, Gu et al. 2013).



*Figure 1.25 Overview of the Infinium Human Methylation 450 BeadChip method. DNA is extracted from the sample and bisulfite converted: bisulphite converts cytosine residues to uracil, but not 5-methylcytosine residues. Bisulfite converted DNA is hybridised onto a Chip and analysed using BeadXpress read. The data are analysed using bioinformatics tools.*

## 1.5 Summary and Aims

Atypical Teratoid Rhabdoid Tumours (ATRTs) and Extra Cranial Rhabdoid Tumours (ECRT) are rare and aggressive malignancies of the central nervous system (CNS). These tumours are most commonly found in infants and toddlers, after the age of 5 years ATRT and ECRT are a rarity. Cytogenetic analysis has revealed deletions or alterations at chromosome locus 22q11, which cause inactivation of *SMARCB1*. These mutation often are acquired *de novo*, however a subset of patients have been show a germ line mosaicism.

Due to the young age of the patients and the rarity of these malignancies there is no defined standard of care, and treatment options have to be tailored to a child's age and metastatic status. Generally treatment involves a combination of therapies including surgery, radiation and high dose chemotherapy. Outcome varies on the basis of age of patient, mutation type and tumour staging. In particular, patients with germ line mutation, younger than 1 year or with metastatic disease have very poor prognosis.

*SMARCB1* has been indicated as a tumour suppressor gene and its role in tumourigenesis has been explored by studies performed in Rhabdoid cell lines. *SMARCB1* suppression is directly linked to disruption of cell cycle by deregulation of the *CDKN2A/ CDK4/Cyclin D1/Rb/E2F* pathway. Furthermore, the re-introduction of *SMARCB1* in Rhabdoid cells is sufficient to induce cell cycle arrest in G<sub>0</sub>-G<sub>1</sub> phase. *In vivo* heterozygous *SMARCB1* knock-out and conditional mice show predisposition to cancer, developing frequent malignancy in the head, neck, the spine and occasionally kidney. Interestingly *SMARCB1*<sup>+/-</sup> are more prone to develop highly aggressive and frequently metastatic tumours than the control mice. The remarkable genomic stability i.e. no other recurrent mutations are found Malignant Rhaboid Tumours and Atypical Teratoid Rhabdoid Tumours suggests that *SMARCB1* deletion is the primary gene for Rhabdoid tumour development. However the discovery of *SMARCA4* mutation in a number of Rhabdoid patients and *SMARCB1* mutation in other CNS tumours indicates a more complex scenario.

*SMARCB1* encodes for SMARCB1 protein which is part of the SWI/SNF chromatin remodelling complexes. SWI/SNF complexes play an important role in epigenetic mechanism, regulating gene expression without alteration in DNA sequence. In fact SWI/SNF complexes directly remodel chromatin structure and recruit a variety of transcription factors, co-activators and co-repressors, modulating gene expression and driving DNA repair. SWI/SNF complexes are important in the differentiation of mammalian tissues (for example neuronal tissue), cell differentiation and in tumourigenesis. The discovery of mutation of a variety of subunits has disclosed the involvement of the SWI/SNF complexes in tumour development. In fact, not only *SMARCB1* mutation but also *SMARCA4*, *ARID1A* and *ARID1B* have been found mutated in a variety of very aggressive malignancies.

Despite the relative simple biology (Jackson, Sievert et al. 2009, Lee, Stewart et al. 2012), there is not an effective cure for Malignant Rhabdoid tumours and Atypical Teratoid Rhabdoid Tumours, mainly because their rarity. Since the discovery of multiple mutations of the SWI/SNF complexes in a variety of tumours, the interest in these malignancies has been raised. Recently Genome-wide approaches (such as SNP-array) have been used to profile Malignant Rhabdoid tumours and Atypical Teratoid Rhabdoid Tumour patients, although lack of patient material and the cost of these techniques have limited their application. Recently microarray profiling of MRT and ATRT patients shows a common gene signature regardless the location when compared with other paediatric tumours (Birks, Donson et al. 2013). Cross-referencing of genome-wide profiling between patients and Rhabdoid cell line identifies up regulation of multiple genes such as *GLI1* (responsible for activation of the Hedgehog-GLI pathway) and *EZH2* (associated with tumour progression and poor prognosis) and down regulation of *IFN*-responsive genes and *Aurora A* (both regulate mitotic spindle checkpoint) (Jagani, Mora-Blanco et al. 2010). However Malignant Rhabdoid Tumours and Atypical Teratoid Rhabdoid Tumours display different gene signatures, implying that *SMARCB1* modulates different factor in base of the tissue.

More recently, high-throughput sequencing of RNA and methylation profiling has been extensively used to identify the epigenetic association with gene expression in

tumourigenesis. This has allowed cancer researchers to characterise tumourigenic events and finally to identify specific genes and pathways involved in tumour development. At this time there are no published data available for RNA-seq or methylation profiling of Malignant Rhabdoid tumours / Atypical Teratoid Rhabdoid Tumours patient or cell lines. Application of NGS and DNA methylation microarrays in studying Rhabdoid tumourigenesis would improve the knowledge of the disease, permitting a better stratification of patients and better therapies.

This project aims to discover the precise mechanism of action of SWI/SNF complexes in Rhabdoid Tumours and the effect of *SMARCB1* loss on function, with the following specific aims.

1. To understand how loss of *SMARCB1* affects epigenetic mechanisms and ultimately gene expression, using novel DNA-methylome profiling and Next Generation Sequencing. Characterisation of the downstream effects of *SMARCB1* loss will aid understanding of the biology of Rhabdoid Tumours as well as the process of chromatin remodelling as a whole.
2. To achieve a comprehensive genome-wide analysis of Rhabdoid malignancies; to investigate differential patterns of DNA methylation and gene expression in primary material that may provide a better diagnosis and stratification of patients.
3. To cross-reference next generation transcriptome and methylome data of primary Malignant Rhabdoid tumours and in functional model primary data to create a genome-wide catalogue of *SMARCB1* –dependent tumourigenic changes.

## **Chapter 2      Materials and Methods**

## 2.1 Primary tumour cohorts

A representative cohort of 39 primary Malignant Rhabdoid Tumours was used for investigations reported in Chapter 4, 6 and 7. The cohort comprised 15 Atypical Rhabdoid Tumours (Table 2.1) and 24 Extra Cranial Rhabdoid Tumours (Table 2.2) were collected from multiple centres throughout the UK in collaboration with the CCLG bio bank ([www.cclg.org.uk/tissue-bank](http://www.cclg.org.uk/tissue-bank)). All samples were evaluated at the referring centre and subsequently centrally review by the CCLG bio bank.

Sample ID	Centre	Location	Age Year.Month	Sex
NMB776	Bristol	Brain	1.5	M
NMB777	London/GOSH	Brain	1.5	F
NMB778	London/GOSH	Right frontal tumour	6.1	F
NMB834	London/GOSH	Frontal tumour	6.6	F
NMB836	Southampton	Posterior fossa	0.6	M
NMB842	Birmingham	Left fronto- parietal	1.5	F
NMB843	Southampton	Intramedullary region	1.2	F
NMB846	Bristol	Frontal recurrence	1.8	M
NMB853	London/GOSH	Brain	1.4	M
NMB854	London/GOSH	Posterior fossa	2.1	M
NMB856	Newcastle	Posterior fossa	4.8	M
NMB876	Bristol	Brain	1.5	M
NMB878	London/GOSH	Brain	1.5	F
NMB885	London/GOSH	Right frontal tumour	6.1	F
NMB888	London/GOSH	Frontal tumour	6.6	F

*Table 2.1 Primary ATRT cohort clinical data. Patient ID, centres where samples were collected age in years at diagnosis, location and sex are shown.*

Sample ID	Centre	Location	Age Year.Month	Sex
NMB835	Bristol	Abdomen	1.4	M
NMB838	London/GOSH	Abdomen	0.4	F
NMB839	London/GOSH	Abdomen	0.1	M
NMB840	London/GOSH	Thoracic cavity	0.4	M
NMB841	Newcastle	Kidney	1.4	F
NMB844	London/GOSH	Kidney	0.3	M
NMB845	London/GOSH	Kidney	1.2	M
NMB847	Sheffield	Kidney	0.6	F
NMB848	Sheffield	Kidney	6.2	M
NMB849	Sheffield	Paraspinal	0.3	F
NMB850	Sheffield	Cervical spine	2.4	M
NMB852	Glasgow	Kidney	0.3	M
NMB860	Manchester	Kidney	0.5	F
NMB861	Manchester	Bladder	7.7	M
NMB862	Manchester	Liver	1.12	M
NMB863	Manchester	Extra-renal	0.7	F
NMB864	Manchester	Buttock	4.7	F
NMB865	Manchester	Kidney	3.7	M
NMB877	London/GOSH	Left Kidney	3.7	M
NMB879	Leeds	Recurrent Kidney	1.1	M
NMB880	Leeds	Liver	1.3	M
NMB881	Manchester	Kidney	2.4	M
NMB886	Southampton	Kidney	0.7	F
NMB896	Glasgow	Abdomen	8.6	M

*Table 2.2 Primary ECRT cohort clinical data. Patient ID, centres where samples were collected age in years at diagnosis, location and sex are shown.*



## **2.2 Cell lines**

### **2.2.1 Cell line cultures**

The role of *SMARCB1* in tumourigenesis was investigated in four different human cell lines; three were Extra Cranial ECRT (G401, A204, and STA-WT1) and one Atypical Teratoid Rhabdoid Tumours (CHLA-266), all with *SMARCB1* deficiency. The 293T cell line was used as host cells for the lentiviral packaging system. The different culture media used are listed in Table 2.3.

Cells were seeded at an appropriate density in T175cm<sup>2</sup> flasks and incubated at 37°C in the presence of 5% carbon dioxide. Cells were passaged upon growth to 80-90% confluence. Cells were treated with trypsin, and then centrifuged to generate a cell pellet. Pellets were re-suspended in an appropriate volume of media before being counted either manually through the use of a haemocytometer or electronically via the TC-10 automatic cell counter (BIO-RAD). The viability was measured by staining an aliquot of the cells with Trypan blue (BIO-RAD).

Cells were stocked by transferring into cryovials in freezing media (10% DMSO, 50% Foetal Calf Serum, and 40% cell media). Initially the vials were placed in a freezing container (Thermo Scientific) at -80°C for 24 hours and subsequently transferred to liquid nitrogen.

### **2.2.2 Proliferation assay**

To assess the proliferative capability of stably transfected cells a colorimetric method CellTiter 96® AQueous (Promega) was used. The reagent provided contains MTS tetrazolium compound (3- (4, 5-dimethylthiazol-2-yl) -5- (3-carboxymethoxyphenyl) -2 - (4-sulfophenyl) -2H-tetrazolium) which is bio-reduced by cells into a coloured

Formazan product. The conversion is due to the production of NADPH or NADH by dehydrogenase enzymes in metabolically active cells.

Briefly every 12 or 24 hours 10µl of MTS solution was added directly to the media. The absorbance of Formazan product (directly proportional to the number living cells) was measured at 490 nm using a FLUOstar Omega (BMG Labtech) plate reader. All measurements were taken in triplicate.

Media	Cell line	Composition
<b>RPMI 10%</b>	G401, A204	1% L-glutamine 1% penicillin/streptomycin 10% foetal bovine RPMI 1640 (Sigma-Aldrich)
<b>RMPI 20%</b>	STA-WT1	1% L-glutamine 1% penicillin/streptomycin 20% foetal bovine 1% sodium pyruvate RPMI 1640 (Sigma-Aldrich)
<b>IMDM 20%</b>	CHLA-266	1% L-glutamine 1% penicillin/streptomycin 20% foetal bovine IMDM (Sigma-Aldrich) 0.1% insulin –transferrin –selenium (Invitrogen)
<b>DMEM 10%</b>	293T	1% L-glutamine 1% penicillin/streptomycin 10% foetal bovine DMEM (Sigma-Aldrich)

*Table 2.3 Media used for culturing cell lines.*

### **2.2.3 Treatment of Malignant Rhabdoid tumours cell lines with 5-aza-2'-deoxycytidine**

The 5-azaCdR treatment was conducted over a period of 5 days. Briefly 5-azaCdR (Sigma-Aldrich) was added at a concentration of 0.1mg/mL. On days 3 and 4, the media in each flask was replaced with fresh media and supplemented with 5-azaCdR where necessary. On day 5, cells were harvested by trypsinisation. Cell pellets were stored at -80 °C.

## **2.3 SDS-PAGE and Western Blot**

### **2.3.1 Protein extraction and quantification**

Protein was extracted from  $1 \times 10^6$  cells. 50  $\mu$ l ice cold urea buffer (90% 8.8 M urea, 2% 5 M  $\text{NaH}_2\text{PO}_4$  and 8% 1 M Tris pH 8.0) was added to  $1 \times 10^6$  cells and vortexed for 15 seconds. Samples were incubated for 30 minutes on ice and then centrifuged at 15,000 rpm for 30 minutes at 4°C. The supernatant was collected and stored at -80°C. Proteins extracted were quantified using the BCA<sup>TM</sup> Protein Assay Kit (Pierce) according to the manufacturer's instructions. Briefly in a 96 well plate 200 $\mu$ l of Working Reagent was added to 5 $\mu$ l of each protein sample. A series of standard protein concentrations containing bovine serum albumin (from 2mg to 0 mg) were prepared using the standard provided in the BCA. The samples were incubated at 37°C for 30 minutes, the absorbance of the protein solutions were measured at 562 nm using a FLUOstar Omega (BMG Labtech) plate reader.

### **2.3.2 Protein electrophoresis and blotting**

Loading buffer (see Appendix) was added to 50  $\mu$ g of protein and the sample denatured at 99°C for 5 min, spun briefly and kept on ice. The samples and a molecular weight marker (Movex Sharp Protein Standard, Invitrogen) were loaded in a 4-20%

polyacrylamide Amersham ECL Gel (GE Healthcare, UK). The run was performed at a constant 160 mV in Tris – Glycine Running Buffer. The electrophoresed proteins were transferred to a nitrocellulose membrane (Pierce) in a Tris-Glycine-Ethanol Transfer Buffer, by application of an electrical field of 100 mV for 45 min. They were then incubated in blocking solution. All solutions are listed in Table 2.1

Buffer	Composition
<b>Running Buffer</b>	0.3% w/v Tris (Sigma Aldrich) 1.44% w/v Glycine (Sigma Aldrich) 0.1% w/v SDS (Sigma Aldrich) dd H <sub>2</sub> O
<b>2X Loading Buffer</b>	2% w/v Bromophenol Blue (Biorad) 5% v/v 1 M Tris, pH 7 (Sigma Aldrich) 25% v/v mL 20% SDS (Sigma Aldrich) 2% v/v β-Mercaptoethanol (Sigma Aldrich) ddH <sub>2</sub> O
<b>Transfer Buffer</b>	0.3 % w/v Tris (Sigma Aldrich) 1.44 % w/v Glycine (Sigma Aldrich) 0.1 % w/v SDS (Sigma Aldrich) 20 % v/v Absolute Ethanol (Sigma Aldrich) dH <sub>2</sub> O
<b>Blocking solution</b>	5% w/v Non-Fat Dry Milk (Sigma Aldrich) 1X TBS 1% v/v Tween-20 (Sigma Aldrich) dH <sub>2</sub> O
<b>TBS pH 7.6</b>	0.24% w/v Tris (Sigma Aldrich) 0.84% w/v NaCl (Sigma Aldrich)
<b>T-TBS</b>	1X TBS (Sigma Aldrich) 1% v/v Tween-20 (Sigma Aldrich)

*Table 2.4 Buffers used for Western Blot*

The membrane was incubated in a solution composed of primary antibodies (in blocking buffer (1:1000) (Brg-1 H-88, Santa Cruz Biotechnology; antibody BAF47, BD Biosciences) for 1 hour at room temperature. It was washed three times in a solution of T-TBS (Table 2.4 ) then incubated in a diluted specific secondary antibody (1:5000) for 1 hour at room temperature and washed in T-TBS as before. The blot was developed using SuperSignal West Pico Chemiluminescent Substrate (Pierce) and imaged using GBox Chemi XL1.4 (Syngene). All antibodies are indicated in Table 2.5.

Antibody	Vendors
<b>Brg-1 H-88 (sc-10760)</b>	Santa Cruz Biotechnology
<b>Anti-BAF47 (612110)</b>	BD Biosciences
<b>Anti-Beta Actin (ab49900)</b>	Abcam
<b>HRP Goat Anti-Mouse IgG (554002)</b>	BD Biosciences
<b>HRP Rabbit Anti-Mouse IgG (ab6728)</b>	Abcam

*Table 2.5 Antibodies used for Western Blot*

## 2.1 Expression quantitation by real time PCR

### 2.1.1 DNA/RNA extraction, purification

DNA was extracted from pellets of  $1 \times 10^6$  cells using the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's instructions. RNA was extracted from pellets of  $1 \times 10^6$  cells using the Total RNA Purification Kit (Norgen Biotek). Briefly the cells were re-suspended in the kit Lysis buffer and DNA or RNA bounded to a membrane. The membrane was washed twice with a kit washing buffer. DNA and RNA were eluted

in an appropriate volume of TE buffer and ddH<sub>2</sub>O (Invitrogen) respectively. DNA samples were stored at -20 °C, and RNA samples at -80 °C.

#### *2.1.1.1 Phenol/chloroform extraction of DNA*

An equal volume of Phenol/Chloroform/Isoamyl alcohol (25:24:1, Sigma Aldrich) was added to the DNA sample. The sample was then centrifuged at 14000 x g for 5 minutes at 4°C. The aqueous phase was transferred into a new tube and an equal volume of Chloroform (Sigma Aldrich) was added. The upper aqueous phase was removed to a new tube after the centrifugation at 14000 x g for 5 minutes at 4°C. 1µl Glycogen (Fermentas) (20mg/ml) and 10% 3M Sodium Acetate (Sigma Aldrich), pH 5.5 was added prior to the addition of two volumes of Absolute Ethanol (Sigma Aldrich). After incubation at -80°C for 30 minutes, the sample was centrifuged at 14000 x g for 10 minutes at 4°C. The supernatant was discarded and the pellet was washed with 1mL 70% ethanol. The pellet was left to air dry then dissolved in an appropriate volume of ddH<sub>2</sub>O or TE.

#### **2.1.2 DNA/RNA quantification and quality assessment.**

Quantification of DNA and RNA was conducted using the NanoDrop 1000 (Thermo Scientific). DNA/RNA purification was performed on samples and DNA and RNA quality has been assessed by Bioanalyser (Agilent Technologies), using Agilent DNA 1000 Kit and Agilent RNA 6000 Nano Kit respectively, following the manufactures procedure. Briefly, the chips were filled with a gel matrix and DNA or RNA ladder was added to each well; in the provided well DNA or RNA sizing ladder has been added. The chip was then vortexed and run on Agilent 2100 Bioanalyser.

### 2.1.3 Oligonucleotides

All oligonucleotides were synthesised by Sigma Aldrich and re-suspended in nuclease free H<sub>2</sub>O (Invitrogen) to a concentration of 100µmol/µl. Primers for q-PCR were designed (Table 2.6) using Primer3Plus software (version 4.0.0 and previously; <http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/>). The transcript sequences of genes were taken from Ensembl (<http://www.ensembl.org/index.html>). The primers were checked for non-specific binding using UCSC In-Silico PCR software (<https://genome.ucsc.edu/>).

Oligonucleotide Name	Length (bp)	T <sub>m</sub> (°C)	GC%	Sequence, 5'-3'
<b>GAPDH Fwd</b>	20	58.4	50	CGACCACTTTGTCAAGCTCA
<b>GAPDH Rev</b>	20	62.5	60.0	GACTGAGTGTGGCAGGGACT
<b>INI1 Fwd</b>	20	59.5	50.0	CAGCCTGGGAGAAAGAGAAA
<b>INI1 Rev</b>	18	61.4	61.1	GCCCCGATGAATGGAGAC
<b>p16 Fwd</b>	20	60.1	55.0	ACCAGAGGCAGTAACCATGC
<b>p16 Rev</b>	20	59.6	60.0	CCTGTAGGACCTTCGGTGAC
<b>p21 Fwd</b>	20	60.7	55.0	CCGAAGTCAGTTCCTTGTGG
<b>p21 Rev</b>	20	61.4	55.0	CATGGGTTCTGACGGACATC
<b>CD44 Ex4 Fwd</b>	19	64.4	57.8	GTCACAGACCTGCCCAATG
<b>CD44 Ex5 Rev</b>	20	64.0	55.0	CCTGAAGTGCTGCTCCTTTC
<b>GLI2 Ex2 Fwd</b>	20	63.2	45.0	AAGCAAGAAGCCAAAAGTGG
<b>GLI2 Ex3 Rev</b>	20	63.9	55.0	TGGTACCTTCCTTCCTGGTG
<b>GLI3 Ex2 Fwd</b>	20	64.2	50.0	CGAACAGATGTGAGCGAGAA
<b>JARID2 Ex7 Fwd</b>	20	64.0	50.0	GTCACGCAGATTCAGCACAT
<b>JARID2 Ex8 Rev</b>	20	64.0	45.0	TAATCAGCCGGAAAAAGCAG
<b>NTN4 Ex8 Fwd</b>	21	63.3	38.0	AAGGCATTCTGTGGAATGAAA
<b>NTN4 Ex9 Rev</b>	20	63.9	50.0	GGATTGGACAAGTGCATCCT

*Table 2.6 q-PCR oligonucleotides used to evaluate differential expression.*

#### 2.1.4 Quantitative polymerase chain reaction

A total of 1000ng of RNA was converted into cDNA using the iScript™ cDNA Synthesis Kit (BIO-RAD) in accordance with the manufacturer's instructions.

The cDNA was amplified using the Platinum SYBR Green qPCR Supermix – UDG with ROX kit (Invitrogen, USA) using a *ViiA7* machine (Applied Biosystems, UK). The master mix provided is composed of AmpliTaq® DNA Polymerase and SYBR® Green cyanine dye, which as a result of binding to DNA absorbs blue light ( $\lambda_{\text{max}} = 497 \text{ nm}$ ) and emits green light ( $\lambda_{\text{max}} = 520 \text{ nm}$ ). General reaction conditions are indicated in Table 2.7.

Polymerase	Reagent	Final Vol.	Cycling
<b>AmpliTaQ®</b>	<b>SYBR® Select Master Mix</b>	5.1µl	50°C, 2'
<b>DNA</b>	Forward Primer (10mM)	0.2µl	95 °C, 10'
<b>Polymerase</b>	Reverse Primer (10mM)	0.2ul	<b>40cycles: 95°C, 15"</b>
<b>( SYBR® Select Master Mix)</b>	DNA Template (1ng/µl)	4ul	<b>60°C, 35"</b>
	dH <sub>2</sub> O	0.5µl	95°C, 15"
			60°C, 1'

Table 2.7 General PCR reaction and conditions

##### 2.1.4.1 Relative quantification of gene expression data using the $\Delta\Delta\text{CT}$ method

The relative quantification measures the relative change in mRNA expression levels of a gene by comparing to the levels of another RNA. This method does not require a calibration curve and the gene expression is compared against a reference gene. In this study we chose *GAPDH* as a reference gene; a constantly expressed endogenous gene.

Gene expression was calculated using the  $\Delta\Delta\text{CT}$  method which assumes that each PCR cycle will double the amount of amplicons in the reaction (amplification efficiency = 100%) or at least that the amplification efficiency is roughly equal between the reference and gene of interest (Livak and Schmittgen, 2001).



The fold change expression values have been calculated using this formula:

$$\text{Fold change expression} = 2^{-\Delta\Delta Ct}$$

$$\text{Where: } \Delta\Delta Ct = [\Delta Ct_{\text{sample1}} - \Delta Ct_{\text{sample2}}]$$

$$\text{And } \Delta Ct = [Ct_{\text{sample}} - Ct_{\text{endogenous control}}]$$

## **2.2 Lentiviral Production and Infection**

For Lentiviral particle production,  $3.5 \times 10^6$  HEK293T cells were co-transfected with 7.5 µg of expression vector and the packaging vectors 6 µg of psPAX2 and 1.8 µg pVSVg using Calcium Phosphate Transfection (CalPhos Transfection Kits, Clontech). Between twelve hours and fifteen hours after transfection, fresh media was added. Forty-eight hours after transfection, supernatant was collected and concentrated by ultrafiltration in Centricon Plus 100 (Millipore). Particles were aliquoted and stored at -80°C.

### **2.2.1 Determining viral titre**

Particle titre was determined by infection of  $5 \times 10^4$  HEK293T cells in serial dilutions, from  $10^{-3}$  to  $10^{-7}$ . Twenty-four hours after infected cells was selected by addition of Puromycin (Invitrogen) to the media to the final concentration of 1 µg/mL. Seven days after infection the cells were fixed with 70% Ethanol (Sigma Aldrich) and stained with Hexadimethrine bromide (Sigma Aldrich). The blue-stained colonies were counted using a microscope at a magnification of 40X. The lentiviral titre was calculated by multiplying the number of colonies per well by the dilution factor.

### 2.2.2 Multiply of infection

Multiplicity of infection (MOI) refers to the number of virions added per cell during infection. MOI was determined by infection of  $5 \times 10^4$  cells of each Rhabdoid cell line in serial dilutions (from MOI of 1 to 100). Twenty-four hours later infected cells was selected by addition of Puromycin (Invitrogen) to the media to the final concentration of  $1 \mu\text{g/mL}$ . Seven days after infection the cells were harvested, counted and the viability was measured by staining an aliquot of the cells with Trypan Blue (BIORAD). RNA was extracted and converted in cDNA to evaluate the relative quantification of gene expression of *CDKN2A*, *CDKN1A* and *CD44* (oligopeptides in Table 2.6). The ideal MOI was that condition presenting the greatest number of infected cells and an overexpression of *CDKN2A*, *CDKN1A* and *CD44*.

### 2.2.3 Stable lentiviral infection of Rhabdoid cells

Lentiviral infection of Rhabdoid cells was carried out aiming to transduce about 60%-80% of the total amount of cells in each experiment. Briefly, 15 minutes before the infection Polybrene (Invitrogen) was added into Rhabdoid cells culture media at a final concentration of  $8 \mu\text{g/mL}$ . Viral particles (MOI of 100) were directly added to the media and the cells were incubated for 24 hours at  $37^\circ\text{C}$  and 5%  $\text{CO}_2$ . To select for infected cells, Puromycin (Invitrogen) was added to the media to a final concentration of  $1 \mu\text{g/mL}$ . The selection was prolonged up to 36 hours (*SMARCA4 D/N*, *SMARCA4 w/t* and *C2979T SMARCA4* stable infection) or 7 days (*SMARCB1* stable infection) post-infection, when cells were harvested.

## 2.3 Bacterial strains

Bacterial strains used in this study were all genetically modified *Escherichia coli* organisms High Efficiency 5-alpha Competent *E. coli* (NEB).

### 2.3.1 Plasmids

#### 2.3.1.1 *PCDH-CMV-MCS-EF1-Puro*

The self-inactivating, CMV-derived, lentiviral vector pCDH-CMV-MCS-EF1-Puro was purchased from Systems BioScience. pCDH-CMV-MCS-EF1-Puro contains Ampicillin resistant genes for selection of the plasmid in *E.coli* and a Puromycin-resistant marker for selection of the transfected/transduced cells.

#### 2.3.1.2 *SMARCB1 wild type*

The pCDH-EF1-PURO-SMARCB1 plasmid was produced by PCR amplifying the whole *SMARCB1* coding sequence from pCDNA 3.1-*SMARCB1* (a gift from Frederique Quignon, Institute Curie). Restriction sites for XbaI and BamHI were added to the Forward and Reverse primers respectively. The PCR product was digested and directionally ligated into the multiple cloning site of pCDH-EF1-Puro (Systems Biosciences).

#### 2.3.1.3 *pBABE SMARCA4 wild type and dominant negative*

pBABE SMARCA4 and pBABE SMARCA4 (K->R) were purchased from Addgene (plasmid 1959 and plasmid 1960). Propagation of all the vectors (excluding pSIN-SIEW) was conducted in the *E. coli* strain DH5 $\alpha$ .

In order to transfect the human cell lines using lentiviral technology, both *SMARCA4* sequences were cloned into the more suitable pCDH-CMV-MCS-EF1-Puro expression vector.

Also the pBABE SMARCA4 wild type was modified in order to create the C2979T mutated *SMARCA4* clone.

#### 2.3.1.4 Packaging plasmids

Two components are required to generate second generation Lentivirus: Packaging Plasmid for producing viral particle and envelope plasmid along with the expression vector. The plasmids used are psPAX2 and pMD2G (kindly donated by Dr Paul Sinclair). The former contains a robust CAG promoter for efficient expression of packaging proteins; the latter plasmid contains the *VSV-G* coding sequence, required for the pseudotyping of the viral envelope.

#### 2.3.2 Bacterial culture

Bacterial strains were grown in Luria-Bertani (LB) (Table 2.8) medium with antibiotic selection from a single colony. Stocks of bacteria were stored at -70° C in a minimum of 25% glycerol (v/v). Microbiological LB agar plates or broth media, 10 ml, were inoculated using a metal loop into LB Plate (Table 2.8) cultures were stored at 4°C wrapped in Parafilm.

Buffer	Composition
<b>Luria-Bertani (LB) medium</b>	1% w/v Tryptone (Sigma Aldrich), 0.5% w/v Yeast extract (Sigma Aldrich), 1% NaCl (Sigma Aldrich), pH 7.5, ddH <sub>2</sub> O.
<b>Luria-Bertani (LB) agar</b>	1% w/v Tryptone (Sigma Aldrich), 0.5% w/v Yeast extract (Sigma Aldrich), 1% NaCl (Sigma Aldrich), 1.5% agar (Sigma Aldrich), pH 7.5, ddH <sub>2</sub> O

*Table 2.8 Buffer used for bacterial culture and composition*

### **2.3.3 Harvesting of bacteria from broth cultures and extraction of nucleic acids**

Bacterial cultures bigger than 1.5 mL were harvested by centrifugation at 4600 rpm for 15 minutes; cultures smaller than 1.5ml were harvested at 13,000 rpm for 3 min.

Plasmids were isolated from cultures using a plasmid purification kit (Qiagen or Invitrogen), according to manufacturer's instructions. Briefly the bacterial pellets were re-suspended in a provided Lysis buffer with RNase and the DNA bound to a membrane. After elution, the DNA was re-suspended first in isopropanol and then in 70% ethanol. The DNA pellets were re-suspended in an appropriate volume of TE buffer.

### **2.3.4 Quantification of nucleic acids by spectroscopy**

DNA extracted from the plasmid was quantified using the ND-1000 (NanoDrop Technologies). Nucleic acids absorb between 260nm and 280 nm. A ratio of ~1.8 between these wavelengths was taken to indicate good sample purity.

### **2.3.5 DNA manipulation techniques**

#### **2.3.5.1 DNA restriction**

New England Biolabs and Promega restriction endonucleases have been used (Table 2.9). Appropriate restriction endonucleases were selected *in silico* using (<http://tools.neb.com/NEBcutter2>). Restriction enzyme digests were performed according to the manufacturer's instructions. Briefly, 10 units of enzyme were used in a reaction mixture containing appropriate buffer, 1µg of DNA and incubated at 37°C for 1-2 hours in thermo-cycler. The enzymes were inactivated by heat for 20 minutes.

Restriction endonuclease	Manufacture
<b><i>EcoR1, EcoRV, Not1, Sal1 and Xba1</i></b>	Promega
<b><i>BsrG1, BstB1, Sap1, Swa1 and Xba1</i></b>	New England Biolabs

*Table 2.9 Restriction endonuclease used for DNA manipulation*

#### *2.3.5.2 Agarose electrophoresis*

Restriction enzyme digests required for further applications were separated by agarose gel electrophoresis. Agarose (Bioline) concentration varied from 0.8% to 1% (w/v) depending on the predicted size of DNA fragments. Agarose was suspended in TBE, boiled and supplemented with Redgel (Biotium) in a concentration of 0.01% (v/v). DNA samples were mixed with DNA loading buffer (Promega). Sizes of DNA sample were evaluated using a DNA marker (Promega) chosen depending on the estimated size of the products (Table 2). A constant voltage was applied across the gel and the DNA size was visualised using Syngene G-box

#### *2.3.5.3 Purification of digested DNA*

The digested plasmid DNA required for further applications were purified using a Gel Extraction kit (GeneJET™ Gel Extraction Kit) according to the manufacturer's protocol. The gel was dissolved with a provided buffer and the DNA bound to a membrane; after cleaning the column with a wash buffer, the DNA was eluted in TE buffer.

#### 2.3.5.4 Oligonucleotides

All oligonucleotides were synthesised and designed as previously described (see section 2.33). The primers used to re-create a mutated SMARCA4 are in Table 2.10. Primers were designed using the transcript sequence from Ensembl (ENST00000344626) and sequence of pBABE-puro (Plasmid 1764, Addgene). Existing primers sequences (SP6, T7 and LNCX) were taken from the corresponding plasmid supplier's website

Oligonucleotide Name	Length (bp)	T <sub>m</sub> (°C)	GC%	Sequence, 5'-3'
<b>C2729T Fwd</b>	16	60.7	69.0	GCCGCCTGCTGCTGAT
<b>C3565T Rev</b>	16	62.5	69.0	CGATGCGGTGGGCTCA
<b>BRG1 Fwd</b>	20	63.7	50.0	AACGCACAGACCTTCAACCT
<b>BRG1 Rev</b>	20	63.8	50.0	CAGCATCTTGTAGGCCATGA
<b>BRG1 SHORT Rev</b>	20	67.9	55.0	CGGATCCGGGACCTGAAATA
<b>pBABE 5'</b>	17	55.4	52.9	CTTTATCCAGCCCTCAC
<b>pBABE 3'</b>	21	60.4	47.6	GGAATGTGTGTCAGTTAGGGT
<b>pBABE 5' Xba1</b>	25	55.4	52.0	GCTCTAGACTTTATCCAGCCCTC AC
<b>pBABE 3' BstB1</b>	30	60.4	50.0	ACCCTAACTGACACACATTCTT TGCAAGCC
<b>C2729T Fwd Sap1</b>	33	58.0	51.5	AAGCTCTTCTATGGGCACACCG CTGCAGAACAA
<b>C2729T Rev Sap1</b>	33	64.0	63.6	AAGCTCTTCTCATCAGCAGCAG GCGGCGGGGTG
<b>BRG1 Fwd BsrG1</b>	20	64.0	50.0	CAGTCAGCGCTTATGGTCAA
<b>LNCX</b>	25	67.4	48.0	AGCTCGTTTAGTGAACCGTCA GATC
<b>SP6</b>	18	42.6	33.3	ATTTAGGTGACACTATAG
<b>T7</b>	20	50.9	40.0	TAATACGACTCACTATAGGG

*Table 2.10 Oligonucleotides used for screening of bacterial colonies and DNA manipulation.*

### 2.3.5.5 Polymerase chain reaction (PCR)

PCR was used to amplify specific regions of DNA. Template DNA for colony screening or molecular cloning was amplified by PCR, using Gene Amp PCR System 9700 thermocycler. For routine colony screening High fidelity PCR PFU polymerase (Agilent), TaqMan® Fast Master (Applied Biosystem) and Phusion® High-Fidelity DNA Polymerase (NEB) were used according to manufacturer's instructions. For negative controls the same reactions were run without the addition of a DNA template. General reaction conditions are indicated in Table 2.11.

Polymerase	Reagent	Final Conc.	Cycling
<b>Fast PCR mix (Applied Biosciences)</b>	GeneAmp® Fast PCR Master Mix (2x)	1x	94°C, 5/10'
	Forward Primer (10mM)	1µM	<b>35 cycles:</b> 94°C, 1'
	Reverse Primer (10mM)	1µM	XX°C, 1'
	DNA Template	XX	72°C, 25"/kb
	dH <sub>2</sub> O up to 20µl	XX	72°C, 7' 4°C ∞
<b>Cloned Pfu (Agilent)</b>	Pfu Buffer (10x)	1x	98°C, 45"
	dNTPs (25mM)	250µM	<b>35 cycles:</b>
	Forward Primer (10mM)	0.5µM	98°C, 45"
	Reverse Primer (10mM)	0.5µM	XX°C, 45"
	Pfu DNA polymerase (2.5 u/µl)	1.25ul	72°C, 60"/kb
	DMSO (100%)	3%	72°C, 10'
	DNA Template	XX	4°C ∞
	dH <sub>2</sub> O up to 25µl	XX	
<b>Phusion High-Fidelity DNA Polymerase (NEB)</b>	HF Buffer (5x)	1x	98°C, 30"
	dNTPs (25mM)	200µM	<b>35 cycles:</b>
	Forward Primer (10mM)	0.5µM	98°C, 10"
	Reverse Primer (10mM)	0.5µM	XX°C, 30"
	DMSO (100%)	3%	72°C, 20"/kb
	Phusion Polymerase (2u/µl)	0.5ul	72°C, 7'
	DNA Template	XX	4°C ∞
	dH <sub>2</sub> O up to 25µl	XX	

XX= Varying amount e.g. annealing temperature, DNA template and dH<sub>2</sub>O

Table 2.11 General PCR reaction conditions for each polymerase.



## 2.3.6 Cloning

### 2.3.6.1 Cloning PCR amplicons

High fidelity PCR Pfu polymerase (Agilent) was used to create the specific sequence of DNA. Cloned *Pfu* DNA Polymerase was used according to the manufacturer's protocol: 1µl cloned *Pfu* DNA polymerase buffer (10×), 1µl dNTP mix (25mM), 0.5µg DNA, 1µl cloned *Pfu* DNA polymerase (2.5u/µl) and up to 10µl dH<sub>2</sub>O. Reaction incubated at 72°C for 30 minutes, reaction terminated by gel extraction.

### 2.3.6.2 Transformation

Plasmids were transformed into High Efficiency 5-alpha Competent *E. coli* (NEB), according to the manufacturer's protocol. 1µl pUC19 (provided) and 5µl of ddH<sub>2</sub>O were used as positive control and negative control respectively. Transformants were selected in LB plates containing 100 µg/ml ampicillin.

### 2.3.6.3 Blunt and cohesive ended ligation

T4 DNA Ligase (Invitrogen) was used according to Table 2.12. The reaction was incubated at 16°C, overnight with no dilution after deactivation, product was stored at -20°C (5µl was used for transformed).

	<b>Cohesive End</b>	<b>Blunt End</b>
<b>T4 DNA Ligase (U)</b>	0.1	1.0
<i>Amount of DNA</i>		
<b>Vector (fmol)</b>	3-30	15-60
<b>Insert (fmol)</b>	9-90	45-180
<b>Vector: Insert ratio</b>	3:1	3:1
<b>Temperature (°C)</b>	23-26	14
<b>Time (h)</b>	≥1	24
<b>Reaction volume (μl)</b>	20	20
Reaction stopped with 1μl 0.5M EDTA, then diluted with 80μl dH <sub>2</sub> O and stored at 4°C.		

*Table 2.12 Conditions used for blunt and cohesive end ligations.*

#### 2.3.6.4 TA Cloning

*SMARCA4* blunt ended PCR fragments products have been incorporate the pGEM T-Easy Vector (Promega). The pGEM T-Easy is a commercial vector with a thymine group added to the break in the plasmid. The Lac-Z operon present in the vector allows a preliminary screening on IPTG/X-gal plates.

Before blunt ended ligation 3' adenosine were added to allow for TA cloning. AmpliTaq Gold® DNA Polymerase (Applied Biosystems) was used according to manufacturer's specification: 2μl PCR Buffer I (10×), 0.4μl dATP (10mM), 1.0μg DNA, 0.5μl AmpliTaq Gold® DNA Polymerase (5.0 u/μl) and up to 20μl dH<sub>2</sub>O. Reaction incubated at 72°C for 20mins, reaction terminated by gel extraction.

Ligation was carried out according to the manufacturer's protocol with a molar ratio of vector: insert of 1:3. Reaction conditions were: 5μl Rapid Ligation Buffer (2×), 1.0μl pGEM®-T Easy Vector (50ng/μl), DNA (2μl control insert DNA for positive control, dH<sub>2</sub>O for the background control), 1μl T4 DNA ligase (3 u/μl) and up to 10μl dH<sub>2</sub>O. If needed reaction was scaled to allow for larger quantities of insert DNA. The reaction was incubated overnight at 4°C and 2μl transformed as above.

#### 2.3.6.5 Site-directed mutagenesis: C2979T *SMARCA4*

To create the C2979 mutation in the *SMARCA4* gene the PCR mutagenesis protocols by Ko and Ma (2005) were followed. Short and long mutant fragments were generated by PCR (Phusion, NEB; 98°C, 30"; 35 cycles: 98°C, 10"; 55°C, 30"; 72°C 30/60"; 72°C, 7'), then products were purified by phenol/chloroform extraction and digested by SapI at 37°C for 2 hours (NEB). The digested fragments were purified by standard phenol/chloroform extraction and ligated with 400U of T4 DNA ligase (NEB) at 16°C overnight to generate a ~3000bp fragment. The ligated fragment was purified from 0.7% agarose gel by Gel Extraction Kit (Fermentas). 1µl was diluted 1:100 (dH<sub>2</sub>O) as a template for a PCR using Phusion® High-Fidelity DNA Polymerase (NEB) with BsrG1\_FOR primer and 3'pBABE untagged primer (product ~3000bp); the insert was ligated into pGEM T-Easy (Promega). Sequence validation used sequencing primer sets 8-14 with the C2729T mutation set (Table 2.13)

#### 2.3.6.6 Sequencing

To confirm that the correct *SMARCA4* and *SMARCB1* sequences were cloned in the various expression vectors and intermediate plasmids, sequencing was performed using the primers listed in Table 2.13. Sequencing was performed by DBS Genomics (Durham) and resultant traces were analysed using ABI Sequence Scanner software (Applied Biosystem).

Oligopeptide Name	Length (bp)	T <sub>m</sub> (°C)	GC%	Sequence, 5'-3'
<b>BRG1 SEQ1 Fwd</b>	18	63.84	66.67	TCCTTCTCTAGGCGCCGG
<b>BRG1 SEQ1 Rev</b>	20	60.11	50.00	TGCATGTTGTCCTGAGGGTA
<b>BRG1 SEQ2 Fwd</b>	18	60.58	66.67	CCACTGTCTGCAGCTCCC
<b>BRG1 SEQ2 Rev</b>	20	58.89	60.00	GGCTGGAAGTGGACTAGAGG
<b>BRG1 SEQ3 Fwd</b>	18	62.73	66.67	CCCAAGGTTACCCCTCGC
<b>BRG1 SEQ3 Rev</b>	20	60.21	50.00	GAGGCCTGCTGTAATTTGGA
<b>BRG1 SEQ4 Fwd</b>	18	63.57	61.11	CGGGATGCAGCAGCAGAT
<b>BRG1 SEQ4 Rev</b>	19	61.20	57.89	TCTGCTTCTGGTGCAGTGG
<b>BRG1 SEQ5 Fwd</b>	20	60.60	55.00	ACCCCTCAGAAGCTGATTCC
<b>BRG1 SEQ5 Rev</b>	20	59.85	55.00	CTCCTGCTCGATCTTCTGCT
<b>BRG1 SEQ6 Fwd</b>	20	60.36	55.00	GGAGACAGCCCTCAATGCTA
<b>BRG1 SEQ6 Rev</b>	20	60.41	50.00	GCGTCTGTCCTTCTGCATTT
<b>BRG1 SEQ7 Fwd</b>	20	60.91	55.00	GCTCATGGCTGAAGATGAGG
<b>BRG1 SEQ7 Rev</b>	20	59.95	45.00	CATCTTGCTTGGCATTCTCA
<b>BRG1 SEQ8 Fwd</b>	20	60.02	55.00	AGCGATGACGTCTCTGAGGT
<b>BRG1 SEQ8 Rev</b>	20	60.06	55.00	TCGTACGTCGTCAGCAAGAC
<b>BRG1 SEQ9 Fwd</b>	20	60.15	55.00	CTCCGTGGTGAAGGTGTCTT
<b>BRG1 SEQ9 Rev</b>	20	59.87	50.00	GCAGACATGTGCACTTGAT
<b>BRG1 SEQ10 Fwd</b>	20	59.98	55.00	CCGACGACTCAAGAAGGAAG
<b>BRG1 SEQ10 Rev</b>	20	59.27	50.00	TGCTGAGCAGGAAGATGAAG
<b>BRG1 SEQ11 Fwd</b>	20	60.66	55.00	CTCAGGCTTGATGGAACCAC
<b>BRG1 SEQ11 Rev</b>	20	60.23	40.00	ATGCGCATGAACAGATCAAA
<b>BRG1 SEQ12 Fwd</b>	20	60.12	55.00	GACGAGACCGTCAACCAGAT
<b>BRG1 SEQ12 Rev</b>	20	60.05	55.00	CCACTGCTGCTGTCTTGTA
<b>BRG1 C2979T Fwd</b>	20	60.07	50.00	GCCAAGATCCGTTGGAAGTA
<b>BRG1 C2979T Rev</b>	19	59.87	52.63	AGGTGCTGCAGCTCTTGAA
<b>PCDH SEQ Rev</b>	18	60.80	67.00	GGAAGTGTGGGCGATGTGC
<b>INI1 SEQ Fwd</b>	20	56.40	45.00	TGACGCCTGAGATGTTTCA
<b>INI1 SEQ Rev</b>	20	58.40	50.00	CCACATGGATGTTTCAGCTTG

*Table 2.13 Oligonucleotides used for sequencing recombinant plasmid.*

## 2.4 Chromatin Immunoprecipitation

Chromatin immunoprecipitation was performed in the A204 and G401 Rhabdoid cell line, using the MAGnify ChIP Kit (Invitrogen). Briefly at 7-days post infection with *SMARCB1* lentivirus particles, the cells were detached by trypsin treatment. Up to  $10^7$  cells were re-suspended in 500  $\mu$ l of PBS and formaldehyde was added for a final concentration of 1% and incubated for 8 min. Room-temperature 1.25 M glycine was added to stop the reaction and incubated for 5 minutes at room temperature. The cells were washed twice with cold PBS and  $1 \times 10^6$  cells were re-suspended in 50  $\mu$ l of provided lysis buffer and incubated on ice for 10 min.

Cell lysate was sheared into ~100–300 bp fragments using Biorupt Standard (High Power, 30 cycles of 30 seconds ON, 30 seconds OFF). The sonicated mixture was spun at 14,000 rpm 15 minutes at 4°C and diluted following the manufacture's direction. 700  $\mu$ l RIPA Buffer were added to the supernatant. Aliquots were stored at –80°C in low binding tubes. The size of the chromatin was checked by agarose gel electrophoresis of purified and de-cross-linked samples using a 2% agarose gel.

The antibodies (Table 2.14) were coupled to Dynabeads® (Invitrogen), incubated for 1 hour at 4°C. Chromatin was bound to the beads by 2-hours incubation at 4°C. The coupled beads were washed according to manufacturer's instructions using the provided washing buffer.

Reverse formaldehyde crosslinking of the chromatin was performed as described in the manual, adding to the beads the de-crosslinking buffer and the Proteinase K (provided in the kit). DNA purification magnetic beads were used to isolate and purify the chromatin.

Antibody	Vendors
<b>Brg-1 H-88 (sc-10760)</b>	Santa Cruz Biotechnology
<b>INI1 C-20 (sc-16189)</b>	Santa Cruz Biotechnology
<b>IgG Mouse</b>	Invitrogen (MAGnify kit)
<b>IgG Rabbit</b>	Invitrogen (MAGnify kit)
<b>IgG Goat (ab37373)</b>	Abcam
<b>H3k27me (c15410069)</b>	Diagenode
<b>H3k4me (c15410030)</b>	Diagenode

*Table 2.14 Antibodies used for Chromatin Immunoprecipitation*

## 2.4.1 Chromatin enrichment determination

Real-time PCR was used to quantify enrichment levels using the purified DNA obtained by ChIP. All qPCR experiments in this study were carried out as described in section 2.3.4. Oligonucleotides used were previously used in ChIP experiments in Rhabdoid cell lines (Kuwahara, Charboneau et al. 2010, Kuwahara, Wei et al. 2013).

### 2.4.1.1 Data analysis and normalisation of ChIP-qPCR data: Percent input method

This method represents the ChIP-qPCR results of the IPs as a percentage of the total amount of chromatin that was present before the immunoprecipitation step.

First, triplicates were averaged. Second, the input Ct value was corrected since the input sample extracted was one hundredth of the total chromatin in the IP samples before immunoselection. A 10-fold increase in product is represented as -3.32.

$$Ct_{\text{Corrected input}} = (Ct_{\text{input}} - 3.32)$$

Then the  $\Delta Ct$  between the IP samples and the input sample was calculated for each of the 3 replicates.

$$\Delta Ct = (Ct_{\text{sample}} - Ct_{\text{Corrected input}})$$

The %input value was calculated as:

$$\%input = 2^{-(\Delta Ct)}$$

The data was then corrected for any background signals indicated in the IgG control:

$$\%input_{corrected} = \%input_{Ab} - \%input_{IgG}$$

## 2.5 *Assessment of gene methylation status*

### 2.5.1 *Assessment of 5-aza-2'deoxyctidine-induced demethylation*

RNA from 5-azaCdR-treated cell lines was first extracted and then converted into cDNA (see section 2.2.1 and 2.2.4). The induced de-methylation by the 5-aza treatments was tested by PCR using the primers outlined in Table 2.15.

Oligonucleotide Name	Length (bp)	Tm (°C)	GC%	Sequence, 5'-3'
<b>DAZL Fwd</b>	24	62.4	41.7	TGAAACTTATATGCAGCCCACAAC
<b>DAZL Rev</b>	24	61.8	45.8	AATAAGCCGGAGGTACAACATAGC
<b>RASSF1A Fwd</b>	21	63.3	57.1	CCTCTGTGGCGACTTCATCTG
<b>RASSF1A Rev</b>	21	61.8	47.6	TAGTGGCAGGTGAACTTGCAA
<b>HTATIP2 Fwd</b>	21	67.4	57.1	CAGCATGGCCGAAACAGAAGC
<b>HTATIP2 Rev</b>	24	67.4	50.0	ACCCAGGCAACAGAATCCAACATC
<b>GAPDH Fwd</b>	23	61.7	43.5	CAAGGTCATCCATGACAACCTTG
<b>GAPDH Rev</b>	22	64.1	59.1	GTCCACCACCCTGTTGCTGTAG

*Table 2.15 Oligonucleotide used for sequencing bisulfite-converted DNA.*

PCR was performed using AmpliTaq Gold DNA Polymerase (Life Technologies) or Phusion High-Fidelity DNA Polymerase (NEB) according to the reaction conditions depicted in Table 2.16. PCR products were separated by gel electrophoresis in a 1% agarose, before being imaged using the SynGene imaging system.

Polymerase	Reagent	Final Conc./Vol.	Cycling
<b>Fast PCR mix (Applied Biosciences)</b>	10x Buffer II	2µl	94°C, 5'
	MgCl <sub>2</sub>	1.2µl	<b>40 cycles:</b>
	dNTPs (5 mM)	200µM	<b>94°C, 30"</b>
	Forward primer (10 µM)	0.5µM	<b>XX°C, 30"</b>
	Reverse primer (10 µM)	0.5µM	<b>72°C, 30"</b>
	Phusion Polymerase (2u/µl)	0.5ul	4°C ∞
	DNA Template	XX	
	dH <sub>2</sub> O up to 25µl	XX	
<b>Phusion High-Fidelity DNA Polymerase (NEB)</b>	HF Buffer (5x)	1x	98°C, 30"
	dNTPs (25mM)	200µM	<b>35 cycles:</b>
	Forward Primer (10mM)	0.5µM	<b>98°C, 10"</b>
	Reverse Primer (10mM)	0.5µM	<b>XX°C, 30"</b>
	DMSO (100%)	3%	<b>72°C, 20"</b>
	Phusion Polymerase (2u/µl)	0.5ul	72°C, 7'
	DNA Template	XX	4°C ∞
	dH <sub>2</sub> O up to 25µl	XX	

*Table 2.16 General PCR reaction conditions for each polymerase used to amplify DNA intended to sequencing.*

## 2.5.2 Bisulphite conversion and sequencing

Bisulfite conversion is used to deaminate unmethylated cytosine to produce uracil in DNA, while methylated cytosines results in protection from the conversion to uracil. DNA bisulfite treatment allows the assessment of DNA methylation status by direct



sequencing. DNA from 5-azaCdR treated cells was bisulfite converted using the EpiTect Bisulfite Kit (Qiagen) in accordance with the manufacturer's instructions.

The human genomic sequences of the candidate genes were extrapolated from the UCSC Genome Browser (<http://genome.ucsc.edu/>). The human genomic sequences of the candidate genes were extrapolated from the UCSC Genome Browser (<http://genome.ucsc.edu/>). Primers specific to the bisulfite-converted sequences of candidate genes were designed using MethPrimer software (<http://www.urogene.org/methprimer/>) and are listed in Table 2.17 and Table 2.18.

PCR was performed using AmpliTaq Gold DNA Polymerase with buffer II and  $MgCl_2$  (Life Technologies) or with Phusion High-Fidelity DNA polymerase (NEB). PCR products were purified using the PureLink PCR Purification Kit (Life Technologies) and sequencing was performed by DBS Genomics (Durham) or Eurofins Genomics (Ebersberg, Germany). The resultant traces were analysed using ABI Sequence Scanner software v2 (Applied BioSystems). The general PCR reaction conditions are delineated in Table 2.18.

Oligonucleotide Name	Length (bp)	T <sub>m</sub> (°C)	GC %	Sequence, 5'-3'
<b>GOLPH1 BDNA RS Rev</b>	26	59.20	42.31	TCCTCCAACAATAAACAAA CATTAAAT
<b>KRT80 Bis Fwd</b>	25	62.90	40.0	TTATTTAAGGGGGTAGGAG AGAAGA
<b>KRT80 Bis Rev</b>	25	62.90	28.00	CCAAAAAACTCTTTAACA ACCAAA
<b>KRT80 BDNA FS Fwd</b>	25	59.89	56.00	TTATTTAAGGGGGTAGGAG AGAAGA
<b>KRT80 BDNA FR Rev</b>	25	57.01	60.00	TACCCAAAATAACTACCCTA CACTC
<b>PSORS1C1 Bis Fwd</b>	30	56.60	6.60	ATGGATTTTATTATATATTTT ATTTTTTTT
<b>PSORS1C1 Bis Rev</b>	27	62.50	37.00	CACCTTCTAAATACCATCCT CTCTAAA
<b>PSORS1C1 BDNA FS Fwd</b>	29	53.82	55.17	TGGATTTTATTATATATTTTA TTTTTTTTT
<b>PSORS1C1 BDNA FS Rev</b>	27	59.94	51.85	TCACCTTCTAAATACCATCC TCTCTAA
<b>PSORS1C1 BDNA RS Fwd</b>	30	52.78	46.67	ATTATTAGTATTTTTTATTTT TTAGATGTT
<b>PSORS1C1 BDNA RS Rev</b>	24	55.36	50.00	TTCTCTCTCTTCTCCAATAA AAAC
<b>S100A10 Bis Fwd</b>	25	58.50	28.00	TTTTGAGTTATTATTTGAAG GAGAG
<b>S100A10 Bis Rev</b>	25	60.40	28.00	CCAAAAAATCAAACCTACA ATAAAC
<b>S100A10 BDNA FS Fwd</b>	23	52.52	52.17	TTGAGTTATTATTTGAAGGA GAG
<b>S100A10 BDNA FS Rev</b>	22	52.13	68.18	AAAATCAAACCTACAATAA ACC
<b>ST5 Bis Fwd</b>	26	62.90	34.60	AAGGAGTTTTTTAGATAGTT GTTGGGA
<b>ST5 BDNA FS Rev</b>	25	53.62	56.00	ACTTCTAAACTTAAAACCA AAATCT
<b>ABR Bis Fwd</b>	24	62.70	33.30	GGTTGGTGTTTGGATTTTTA GATT
<b>ABR Bis Rev</b>	25	62.00	36.00	CAAACCCCTAACTTTCTACC AAATA

*Table 2.17 Oligonucleotides used for sequencing bisulfite-converted DNA.*

Oligonucleotides Name	Length (bp)	T <sub>m</sub> (°C)	GC %	Sequence, 5'-3'
<b>BHLHE40 Bis Fwd</b>	25	61.50	36.00	TAGTTGAGTTAGTTGGTGAT TTGGT
<b>BHLHE40 Bis Rev</b>	25	62.80	32.00	CACAAAAAACTAATATTCC CCTCAA
<b>BMF Bis Fwd</b>	30	58.40	20.00	GGTAGTAAAAATTTGATTTT ATTTAGTGTA
<b>BMF Bis Rev</b>	30	56.70	6.60	ATTCTTTATTTTTATTTTATTT TCAAATAT
<b>BMF BDNA FS Fwd</b>	26	50.68	34.62	AAATTTGATTTTATTTAGTG TAAATA
<b>BMF BDNA FS Rev</b>	30	51.95	26.67	CAAAACATAATATATATTCTT TATTTTAT
<b>BMF BDNA RS Fwd</b>	30	58.01	50.00	CAACAAAAATCTAATTTTCAT CCAATAC
<b>BMF BDNA RS Rev</b>	27	58.50	40.74	TTAGAAAAATGTATTTGTAG TTGAGGG
<b>CCR7 Bis Fwd</b>	27	61.30	29.60	TTAGAAAAATGTATTTGTAG TTGAGGG
<b>CCR7 Bis Rev</b>	26	57.20	23.00	ATAACCAAAAAACAATACC AATATAC
<b>CCR7 BDNA FS Fwd</b>	27	58.35	44.44	TTAGAAAAATGTATTTGTAG TTGAGGG
<b>CCR7 BDNA FS Rev</b>	25	53.82	68.00	TAACCAAAAAACAATACCA ATATAC
<b>CCR7 BDNA RS Fwd</b>	22	55.33	68.18	GTTAGGGGGTAATGTTAGT GTG
<b>CCR7 BDNA RS Rev</b>	30	55.00	56.67	TAACCTAACTACCTTTTACA TAACTACTCT
<b>GALNT1 Bis Fwd</b>	25	62.40	32.00	TATTTGGGTTTTTTGTTAGGT TGAGT
<b>GALNT1 Bis Rev</b>	25	62.50	40.00	ATAACCCTCCTACCTACATT CCATT
<b>GOLPH1 Bis Fwd</b>	24	60.60	20.80	ATTTGTTTGTTTTTTTGTTTT GTG
<b>GOLPH1 Bis Rev</b>	25	58.00	24.0	CAAAAAACTACTTAAAAAA TTCCAC
<b>GOLPH1 BDNA RS Fwd</b>	25	58.43	64.0	TTAGGGATAGTGGTTATGGG TAGTT
<b>ST5 Bis Rev</b>	29	55.7	17.20	AAAAATAAATCTACTTCTAA ACTTAAAC
<b>ST5 BDNA FS Fwd</b>	30	55.81	46.67	GTAAAGAGTTTAAGGAGTT TTAGATAGTT

*Table 2.18 Oligonucleotides used for sequencing bisulfite-converted DNA.*

## **Chapter 3      Profiling and Bioinformatics Strategy**

### 3.1 Introduction

Next generation sequencing has become an important part of molecular biological and patient profiling. In particular, techniques such as RNA-seq, ChIP-seq and methylation arrays are used to give a more systematic analysis of the whole genome, allowing better understanding of epigenetic and regulatory morphology, gene expression regulation and DNA-protein binding.

This project was focused on changes in gene expression and methylation and how the DNA binding of SWI/SNF complex changes after *SMARCB1* re-expression. The strategy aims to select a set of candidate genes/pathways with interesting features as potential Malignant Rhabdoid Tumours drug targets. Furthermore, the experiments carried out in the Rhabdoid cell lines have been cross-referenced to a set of primary data, to characterise the key biological features of these tumours.

### 3.2 RNA sequencing pipeline

The next generation sequencing technologies present more advantages than previous technologies such as microarrays. For instance, RNA-seq does not require a fully annotated transcriptome and has better resolution and sensitivity (Martin and Wang 2011). RNA-seq is more widely used next generation sequencing technique, providing access to high-level transcriptomic information, such as novel promoters and isoforms, fusion genes, aberrations, alternative splicing and novel transcript (Oshlack, Robinson et al. 2010). Many approaches have recently been developed to extract such information from RNA-seq data, however creating a solid analysis pipeline remains an ongoing challenge. Additionally, RNA-sequencing generates millions of raw sequence reads which are processed by multiple computational processing steps. In this study pipeline analysis was tailored to characterise changes in gene expression following *SMARCB1* re-expression in Malignant Rhabdoid Tumour cells and expression patterns in patient material.

### **3.2.1 Samples preparation**

Samples were prepared as indicated in section 1.1.1. After quantification with picogreen and quality evaluation by Agilent Bioanalyser (section 2.1.1), RNA samples were shipped to Aros, Denmark on dry ice for library preparation.

### **3.2.2 Library preparation**

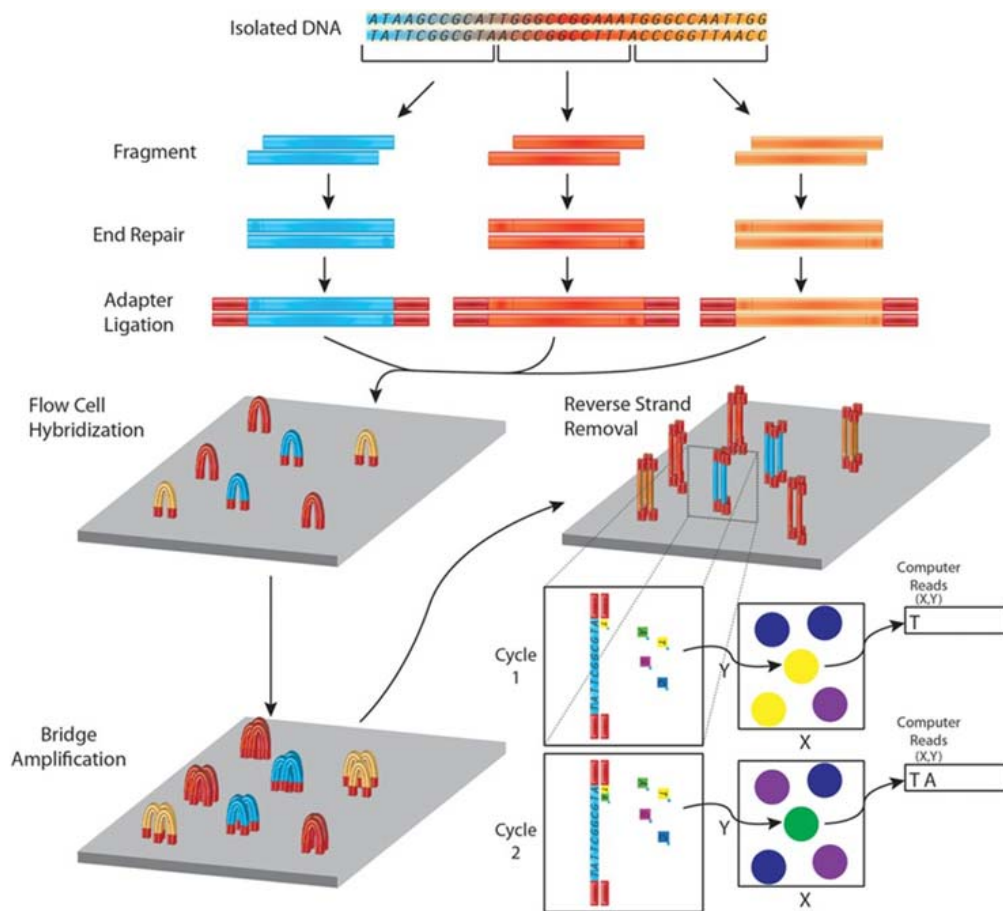
RNA (0.5 – 10 ng) was amplified using the Ovation RNA-Seq system (NuGEN). Briefly cDNA/mRNA hybrid molecules were produced and fragmented to synthesise double-stranded cDNA. cDNA was subject to SPIA (Single Primer Isothermal Amplification) amplification and ends were repaired using T4 DNA polymerase and Klenow polymerase. Single A bases were incorporated to the nucleotide using Klenow 3' and 5' exopolymerase, followed by ligation with a mix of adaptor oligonucleotides using T4 DNA ligase. The adaptor ligated nucleotide of approximately 200 bp was isolated on a 2% agarose gel. The ligated oligonucleotides were amplified for cluster generation by 10 to 15 cycles of ligation mediated PCR (LM-PCR). Subsequently the samples were purified and quantitated using the Agilent Bioanalyser. 2 – 4 pM of the resulting cDNA libraries were used for sequencing.

### **3.2.3 Solexa sequencing (Illumina Hi-seq 2000)**

The Illumina sequencing technology is based on “sequencing-by-synthesis” (SBS) method, which implies the use of DNA polymerase to extend many DNA strands in parallel (Liu, Li et al. 2012). This strategy is characterised by a real-time sequencing without template amplification, using indicators of nucleotide incorporation. An overview of the workflow of Illumina sequencing is given in Figure 3.1.

The flow cell surface is coated with single stranded oligonucleotides that correspond to the sequences of the adaptors ligated during the sample preparation stage. The single

stranded oligonucleotides (corresponding to the library adapter ligated sequences) were bound to the surface of flow cells, and exposed to reagents for polymerase-based extension. The insertion of unlabelled nucleotide and *bst* polymerase at the free/distal end of a ligated fragment "bridges" to a complementary oligonucleotide on the surface. Repeated cycles of denaturation and extension result in an amplified cluster of single molecules in unique locations on the flow cell surface (cluster station). To identify the physical location of a cluster a high resolution image of the flow cell was taken after the first incorporation of fluorescent nucleotide. Each fluorescent nucleotide has reversible terminator chemistry such that one nucleotide may be added followed by flow cell scanning then the terminator chemistry reversed to allow the next fluorescent nucleotide to be added. This process is repeated for every nucleotide sequenced. Every single image of the flow cell taken in each cycle represents an individual base extension at a specific cluster. We aimed to obtain ~90 million reads of ~100 bases, taken from both ends of each cDNA fragment (paired-end reads).



*Figure 3.1 Overview of DNA sequencing using the Illumina platform. RNA is first converted in cDNA and fragmented into smaller fragments, using nebulization. After end repair, specific adapters are ligated to the ends of the fragments, allowing hybridization to a flow cell. Bridge amplification creates clusters of fragments with the very same sequence. One strand of DNA is detached and the clusters are incubated with fluorescently labelled nucleotides. A single base is incorporated at every cycle by a modified DNA polymerase. Sequentially images of the flow cell are recorded and a computer processes which nucleotide was incorporated at each cluster's co-ordinates. Subsequently, images are converted into sequence files by image analysis, base calling, and sequence filtering (Churko, Mantalas et al. 2013).*



### 3.2.4 Primary data processing

The Illumina Genome Analyser (GA), using Sequencing Control Software (SCS) v2.9, produced images files as primary output from a sequencing run. Specifically, the initial image analysis was performed by Genome Analyser Control Software (GCS), followed Real Time Analysis (RTA) for base calling and quality testing. This process (Figure 3.2) includes:

1. Image analysis, which provides the cluster location, cluster intensity and noise estimation.
2. Base calling, which provides the sequencing of bases for each cluster, a confidence level for each base and read filtering, using Bustard RTA software; the results are stored in \*.bcl binary file containing base call and quality for single base in each cycle.
3. FASTQ generation; to continue the analysis \*.bcl files are converted in \*.fastq.gz files (compressed FASTQ files) by CASAVA software, creating a text file containing nucleotide sequence and its corresponding quality scores.

#### 3.2.4.1 Raw data quality control

The raw data were subject to a quality control using FastQC program (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>). The program produces a complete report detailing the composition and the quality of sequence, focusing on insert size distribution, true peak calling, detection of sequence duplications and kmer content. The Fast QC analysis scrutinises each read, scoring according to the Qphred score (Mbandi, Hesse et al. 2014). The Qphred score corresponds to the probability of incorrect base call (Table 3.1). Usually a Qphred 20-30 is set as a threshold to distinguish useful sequence data. In this study the Qphred value threshold was > 30 the quality of sequenced nucleotides generally decreasing slightly from 5' to 3' end.

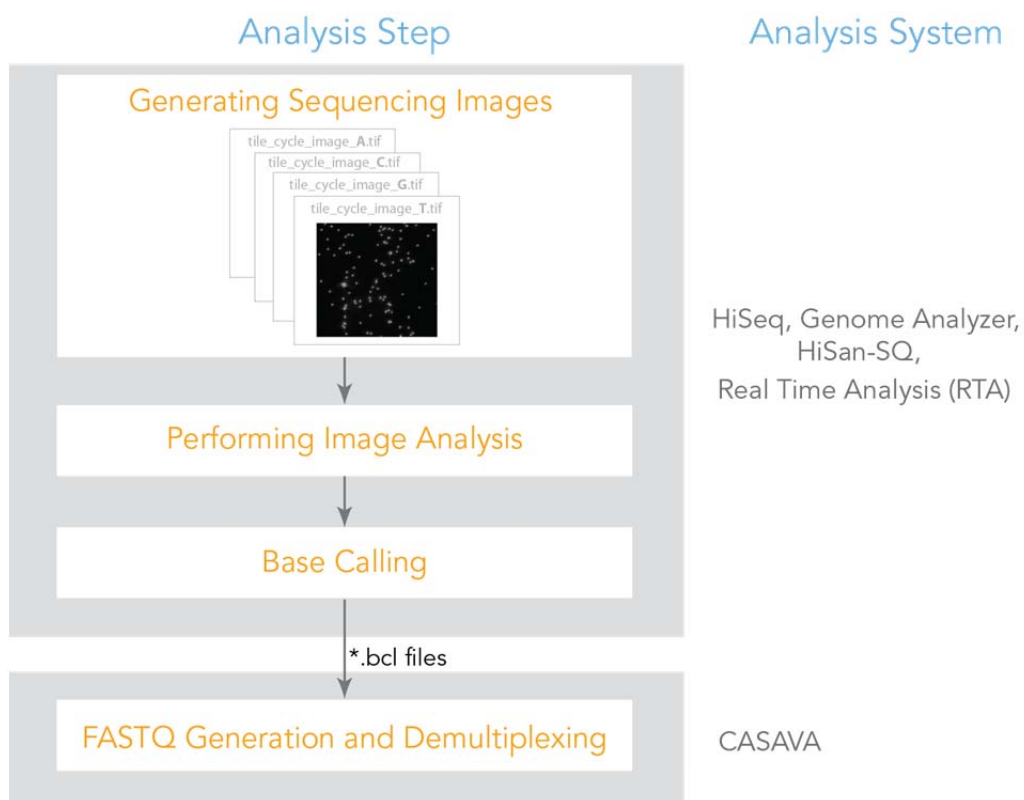


Figure 3.2 Overview of the primary data analysis. The high quality images of the flow cell are processed to extract cluster location, cluster intensity, and noise estimates. In Real Time Analysis (RTA) the processed trace is translated into a sequence of bases, creating a \*.bcl text file. \*.bcl files are converted on FASTQ file using CASAVA software; in this step de-multiplexing is performed.

QPhred quality score	Probability of an incorrect base call	Base call accuracy
10	1 in 10	90%
20	1 in 100	99%
30	1 in 1000	99.9%
40	1 in 10000	99.99%

Table 3.1. Qphred quality score values and accuracy

### *Sequence Alignment (Star Aligner) RNA-STAR: ultrafast universal RNA-seq aligner*

Reads were aligned to the HG19 (UCSC) Homo Sapiens genome using Splice Transcripts Alignment and Reconstruction tool (STAR) (Dobin, Davis et al. 2013). This software permits a novel alignment strategy for *de novo* detection of the splice junctions (exons), not requiring any previous *a priori* annotation. The algorithm is strategically designed to identify maximum mappable length of a read, multiple loci to which a read can be aligned, to estimate the relative probabilities of these alignments. STAR stores alignment of nucleotide sequences in \*.Aligned.out.sam files which are subsequently converted in \*.BAM files to be used in other programs.

#### *3.2.4.2 Alignment Quality Control (SeqQC)*

The quality of the alignment was assessed using SeqQC V2.1 a program designed to check a wide range of metrics showing the overall quality of alignment i.e. number of reads aligned (<http://genotypic.co.in/SeqQC.html>). The program gives an automated overview of sequence files with graphical summary reports including number of sequences, length distribution, low complexity regions, GC content and user specified sequence searches. This is a quality check of raw sequence data, important prior to in-depth analysis HTSeq: high-throughput sequencing analysis.

The HTSeq-count python script was used to quantify the number of reads mapping to exons for each gene (Anders, Pyl et al. 2014). HTSeq analyses how many reads overlap to a given list of genomic features, in this case exon structure of the genes. Additionally to investigate alternative splicing, each exon has been considered as a single feature. In addition HTSeq allows analysing reads that overlap more than one genomic feature. The script permits analysis in three different modes, with respect to overlap resolution: union, intersection-strict, and intersection-nonempty (Figure 3.3). In this study the union mode was chosen and the reads counted according to the GENCODEv17 transcriptome library. Using HTSeq it is possible to obtain a table with counts overlapped with annotated features and discarding multi-alignment reads.

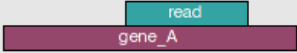
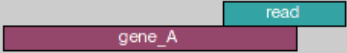




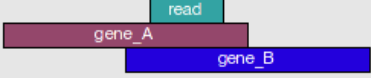
	union	intersection_strict	intersection_nonempty
	gene_A	gene_A	gene_A
	gene_A	no_feature	gene_A
	gene_A	no_feature	gene_A
	gene_A	gene_A	gene_A
	gene_A	gene_A	gene_A
	ambiguous	gene_A	gene_A
	ambiguous	ambiguous	ambiguous

Figure 3.3 Illustration of the three overlaps resolution modes, used in the high-throughput sequencing analysis (<http://sfg.stanford.edu/>).

### 3.2.5 Differential Expression (DEseq)

DESeq is a commonly used R package to test for differential expression analysis of RNA-seq data (Anders and Huber 2010). DESeq is based on the hypothesis of non-linear relationship between the variance and mean expression levels. DESeq performs statistical testing on raw data set obtained from the alignment resulting in differential gene expression analysis, throughout the dynamic range of the data. Dispersions are estimated using expected mean values from the maximum likelihood estimate of log2 fold changes, and optimising the Cox-Reid adjusted likelihood. A generalised linear model is then fitted and tested for significance. The steps are 1) estimation of size

factors 2) estimation of dispersion, 3) negative binomial GLM fitting and Wald statistics. In this study DESeq has been applied to RNAseq count data to search for differential abundance (or differential expression DE) of a particular gene between different conditions (with or without *SMARCB1* re-expression) or different tumour types (Medulloblastoma vs Malignant Rhabdoid Tumours; Atypical Teratoid Rhabdoid Tumours vs Extra Cranial Rhabdoid Tumours). The detection of DE genes is based upon a negative binomial parametric model to test a generalised linear model as a means of hypothesis testing.

Initially the raw data is processed to estimate the effective library size (to calculate the relative size of each library from the count data) and the counts normalised to the effective library size e.g. a library of 120 million reads would have its gene counts reduced by a factor of two when compared to a library of size 60 million. Secondly, DEseq extrapolates the dispersion (a measure of biological and technical variance) directly from the mean. This data may additionally be subjected to a variance stabilising transformation (VST) converting the data into an approximately homoscedastic measure of expression (variance becomes independent of the mean) applicable as an input for a variety of follow up statistical analysis (e.g. Limma, NMF, etc.).

Statistical tests have been used to determine if the estimated changes in expression between the two conditions chosen were significant. DESeq calculates differential expression attributes such as False Discovery Rate (FDR), fold change values and p-values from count data. A variation of Fisher's exact test was used to calculating exact p-values computed from the derived probabilities. To control the false discovery rate, a multiple testing correcting Benjamini-Hochberg procedure was applied to the data (p value adjusted).

### **3.2.6 Differential Exon Usage (DEX-seq)**

DEXseq is a similar tool to DESeq, but it identifies differential splicing or differential exon usage between two sets of conditions (Anders, Reyes et al. 2012). DEXSeq works

by dividing the reads mapped to the same gene into counting bins depending on the exons they overlap. Counting bins do not completely correspond to exons since different transcripts may have different exon sizes in the same genomic location. In that way, a transcript that has varying lengths for the same exon gets assigned to two counting bins, one for the shortest exon, and one for the additional bases of the longer one, plus more counting bins for the remaining exons. Each unit of non-overlapping exon is referred to as an exonic part. DEXSeq also uses a generalised linear model (GLM) to model the read counts in exactly the same way as the DESeq (see above) however dispersion estimations are performed exon by exon instead of gene by gene prior to statistical testing exons are also normalised to the overall number of counts for the whole gene so as to provide a measure of exon usage relative to the expression of the gene.

The analysis of the reads with DEXseq gives values such as conditional-maximum-likelihood estimate for the dispersion, dispersion value used in the test, the p value from the test for differential exon usage, the Benjamini-Hochberg-adjusted p value, and log2 fold change between treatment controls. Furthermore, DEXSeq creates a HTML page for each gene that contains the test results for an interactive browsing of the results.

### **3.2.7 Transcriptome assembly: differential CDS, promoter usage, TSS usage and novel isoforms**

Transcriptome assembly is a qualitative analysis of the genome focused on alternative splicing, antisense expression, extragenic expression, alternative 5' and 3' usage and detection of fusion transcripts. Short reads can be aligned not only to a reference genome but to an independent genome (*de novo* approach). *De novo* assembly aims to merge all possible overlapping sequences to form contigs (Garber, Grabherr et al. 2011). In RNAseq contigs (or transfrags) represent larger fragments of combined and assembled reads, each correspond to transcripts derived from protein-coding genes (Figure 3.4).

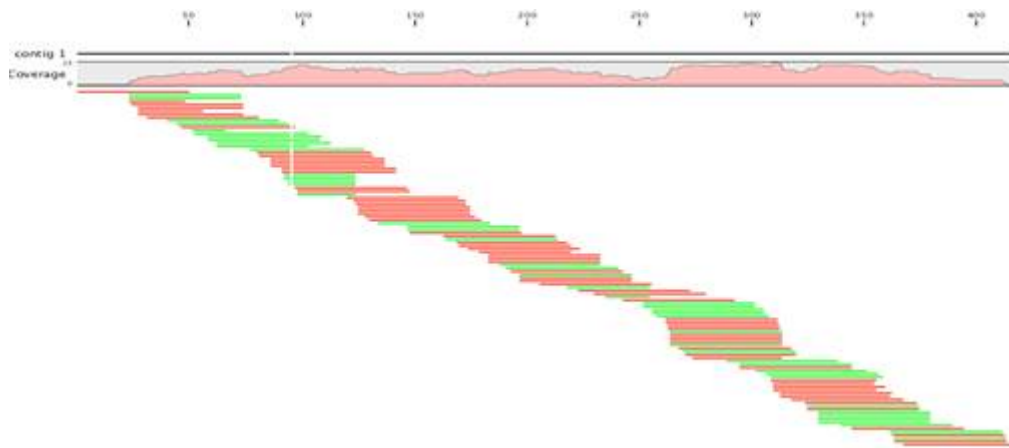


Figure 3.4 Example of a contig assembled by the joining of many short reads (<http://sfg.stanford.edu/>).

In *De novo* assembly analysis lengths and numbers of contigs are usually shorter and more numerous in comparison to the predicted number of transcripts of the reference genome. This is caused by the impossibility of joining all contigs unless there is enough coverage and overlap between reads.

The transcript assembly was performed using Cufflinks, a program which maps the transcript fragments to a list of input data. Cufflinks analyses \*.SAM alignment files and can give an indicative value of the relative abundances of transcripts. Cufflinks calculates the indicative value of the relative abundances based on the number reads of each transcript type. Moreover, it provides coordinates of transcripts and the position of exons in relation to their position on the reference genome. Cufflinks analysis finally provides FPKM values (Fragments Per Kilo-base of exon per Million fragments mapped), which are used to normalise gene expression levels. It also filters out certain artefacts and background using rigorous statistical calculation. The analysis resulted in three different output data set in \*.GTF file format:

1. Gene expression file which contains estimated values of gene-level expression.
2. Transcript expression file which contains estimated values of isoform-level expression.

3. Assembled transcript data set which contains the location of each individually assembled transcript, their FPKM and the exon position.

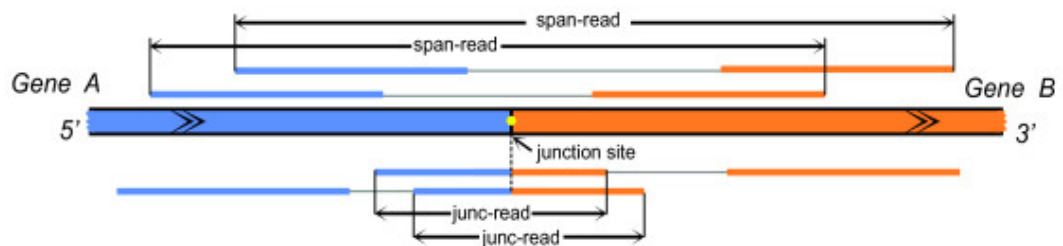
After having individually assembled each alignment, assemblies from each sample were merged together, using the meta-assembler tool Cuffmerge. (Roberts, Pimentel et al. 2011). This strategy has been chosen to avoid assembling the transcript incorrectly; in fact pooling samples in a set can lead to a complex mixture of splice isoforms, increasing the possibility of an incorrect transcript assembly. Cuffmerge was used to merge *de novo* assemblies with the GENCODE v17 transcriptome library and thus identifies where a transfrag is already known (i.e. completely encapsulated within an existing GENCODE v17 gene, a novel isoform (i.e. partially overlaps with an existing GENCODE v17 gene) or a completely novel transcript (i.e. no overlap with an existing GENCODE v17 gene). Once a merged transcript assembly has been produced this may be used as an analysis template for Cuffdiff, which uses statistical methods to calculate differential expression between samples. Cuffdiff is a program developed to reduce biological variability; learning to count the reads for each gene across the replicates Cuffdiff is able to estimate variance between replicates and therefore calculate the significance of the expression changing. Moreover, Cuffdiff can identify other features beyond change in gene expression, such as the presence of differentially spliced genes or genes regulated by promoter switching. Notably, Cuffdiff can identify multiple splice variants of the same gene and evaluate relative abundances of isoforms, using Jensen-Shannon divergence. Gene and transcript expression level changes are listed as simple tabular output, containing fold change, *p*-values, adjusted *p*-values, gene and transcript names and location in the genome. The tool also groups isoforms by TSS (Transcriptional Start Site), searching for changes in relative abundance between conditions. Similarly coding sequences (CDS) are investigated to distinguish gene isoforms with differential expressed protein-coding sequences. Finally, the analyses are visualised using CummRbund, (R package) which helps to manage and visualise the data.



### 3.2.8 Transcript fusion detection

RNA-Seq has been also used as an approach to gene fusion discovery (Maher, Kumar-Sinha et al. 2009, Edgren, Murumagi et al. 2011) and specifically identifies which fusion events are responsible for aberrant transcript (Jia, Qiu et al. 2013).

I use SOAPfuse (Jia, Qiu et al. 2013) in order to detect novel transcript fusions. Initially the paired end reads are aligned to the human genome reference and filtered. The algorithm searches for paired end reads that map in a discordant manner to the genome (span reads) and possible junction reads (junc reads) that can confirm possible fusions (Figure 3.5). SOAPfuse also applies filter and quality controls to characterise and discard artefacts. The filters reject pseudo-genes and genes with extreme homogeneity which are prone to misaligning from further analysis.



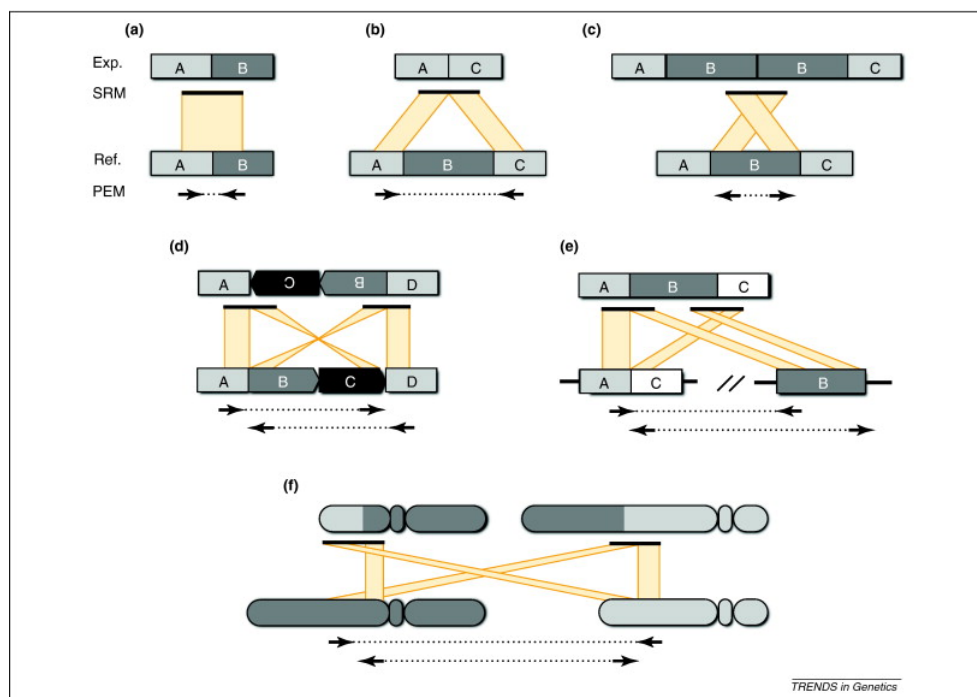
*Figure 3.5 Example of fusion events between gene A and gene B. The fusion event is suggested by discordant alignments to the genome (span reads) and junctions reads (junc reads) (Jia, Qiu et al. 2013).*

After having retrieved pair ends that support possible fusion genes, SOAPfuse is able to determinate the upstream and the downstream gene of the fusion. In particular the determination is made base on the direction of alignment of the span reads against the genome reference (Figure 3.6).

With SOAPfuse it is possible to identify:

1. Interchromosomal and intrachromosomal fusion
2. Gene order inversion

Additionally the algorithm is able to identify the upstream and downstream gene in the fusion events (Figure 3.6). This is possible since the RNAseq data are composed by paired end reads (one starts from the 3', the other from the 5' ends).



*Figure 3.6. Patterns observed for paired-end fusion analysis. a) perfectly matched to the reference genome; b) gene deletion; c) tandem duplication; d) inversion; e) transposon insertion; f) reciprocal translocation. Paired-end read aligned to the reference genome, variation of the distance and orientation of the reads indicate a rearrangement has occurred. Reads that map to the plus strand are indicated as right-facing arrows, those that map to the negative strand as left-facing arrows (Quinlan and Hall 2012).*

### **3.2.9 Identification of viral integration on human genome**

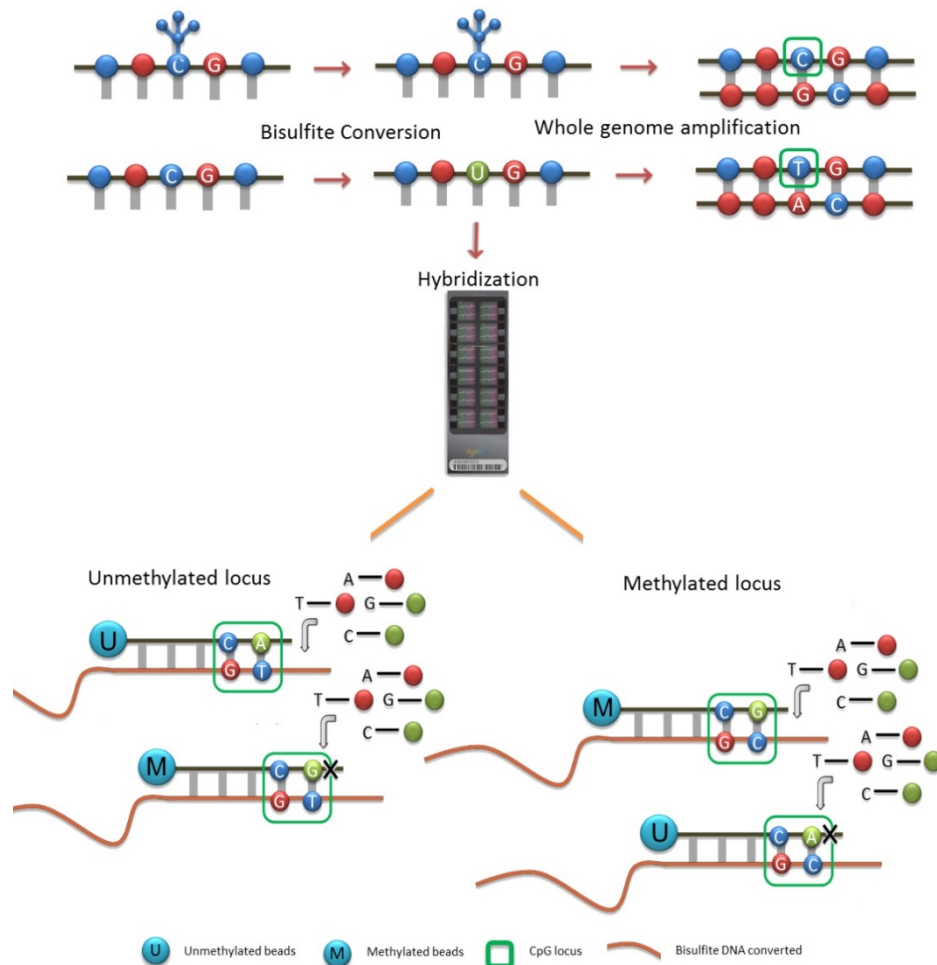
Data nonaligned in the previous alignment step have been analysed for possible virus integration site. In particular, the data have been aligned to a virus database. VirusFinder 2 software have been used to characterise possible viral integration site or fusion (Wang, Jia et al. 2013). Specifically, the tool can detect virus infection, co-infection with multiple viruses, virus integration sites in host genomes. VirusFinder 2 operates in a similar manner of SOAPfuse, mapping contigs to UCSC hg19 human reference and any then aligning any remaining unmapped reads to a viral genome database.

### **3.3 450K Chip Analysis Methylation Pipeline**

Recent advances in genomics and biochemistry have partially explained the functions and mechanism behind DNA modifications. DNA methylation was first observed in bacteria and associated with a series of restriction-modification processes (Bestor 1990). In the last decade the study of methylation events has highlighted the association of DNA methylation with cancer development (Esteller 2008). Different approaches were developed, based upon techniques such as bisulfite conversion (Xi and Li 2009), digestion with methylation-sensitive restriction enzymes (Hashimoto, Kokubun et al. 2007), and affinity purification of methylated DNA (Robinson, Stirzaker et al. 2010). These techniques require an extensive data processing and analysis to extrapolate the information about the methylation status. In this study we genotyped sodium bisulfite treated DNA using the Illumina 450K Human Methylation BeadChip assay to evaluate the genome-wide level of methylation in patients and changes in methylation pattern following *SMARCB1* re-expression in Malignant Rhabdoid Tumour cells.

#### **3.3.1 HumanMethylation450 BeadChip assays**

HumanMethylation450 BeadChip (450K) assays are based upon genotyping bisulfite-converted genomic DNA (gDNA). 450K-array platforms determine a quantitative measure of methylation at CpG by hybridisation of amplified bisulfite-converted genomic DNA to microarrays using methylation-specific oligonucleotides. The 450K DNA Methylation microarray contains 485,512 probes, exploring 19,755 unique CpG islands, as well as 3091 probes at non-CpG sites coverage. This allows coverage to some extent of a total of 21,231 (99%) UCSC RefGenes (Morris and Beck). An overview workflow of the 450K methylation array is given in Figure 3.7.



*Figure 3.7 Infinium Human Methylation 450K BeadChip workflow. Unmethylated cytosines converted to uracil by bisulfite treatment, while 5-methylcytosines are resistant to conversion. The bisulfite converted DNA is subjected to whole genome amplification by PCR with random primers; subsequently the DNA is enzymatically fragmented and hybridized to a chip. The DNA is bound to two different types of beads, one selective for methylated locus (M) and one for the unmethylated (U). CpG targets hybridised with beads are incubated with specific probes which enable single-base extension and detection. On the left an unmethylated locus has been bound with methylated and unmethylated beads. In this case there is a single-base mismatch to the M probe which inhibits extension. If the CpG locus of interest is methylated, the reverse occurs.*

Briefly bisulfite treatment results in conversion of non-methylated Cytosines (C) to Uracils (U) while methylated Cytosine residues are retained as Cytosine. PCR amplification with primers specific for bisulfite modified DNA leads to amplicons where Thymines (T) replace Uracils (U) while the non-modified Cytosines do not change. The amplified and converted oligonucleotides are hybridised to locus-specific oligomer probes bound to individual beads on a BeadChip.

The 450K BeadChip consists of two types of beads. Infinium assay I uses two different site-specific probes, one specific for the methylated locus (M bead type) and the second for the unmethylated locus (U bead type). Subsequently, fluorescent-labelled single base extension provides the array signals. Methylation status is acquired from the fluorescence intensity ratios between unmethylated and methylated bead type (Illumina). Infinium II probes instead are attached to unique beads and use different dye colours (green and red) to differentiate between Methylated and Unmethylated signals, respectively. Dual channel single-nucleotide primer extension is performed, adding single base extension with dideoxynucleotides labels. In this case the methylation state is determined at the single base extension step.

### **3.3.2 Primary data processing**

Raw data representing the fluorescence are Initially processed using Beadstudio v3.2 (Illumina). Specifically, the program converts \*.TIFF files produced by BeadScan software (Illumina) into idat files containing information of paired probes-colour channels. Idat files are processed using the Bioconductor package Minfi (Aryee, Jaffe et al. 2014). It also assess quality of the array and calculates background levels computed from a set of negative controls (Wu, Irizarry et al. 2004). Minfi calculates ratios between methylated (M) and unmethylated fluorescent signals (Beta value) which indicates the quantitative measurement of DNA methylation status. Beta-value is calculated according to the following formula (Martin-Subero, Ammerpohl et al. 2009):

$$\beta = \frac{\text{Max (M intensity, 0)}}{\text{Max (U intensity, 0) + Max (M intensity, 0) + 100}}$$

Max (M, 0) and Max (U, 0) indicate the maximum intensity value between M and 0, U and 0, respectively. The number 100 is a constant to standardize beta values. Beta values are quantitative measurement of DNA methylation status of specific CpGs: the values range from 0 to 1; completely unmethylated and completely methylated respectively.

### 3.3.2.1 Methylation Microarray Quality Control

Beadstudio also produces quality control (QC) files as \*.txt files which are processed with the R package Beadarray (Dunning, Smith et al. 2007) for the identification of spatial artefacts. Firstly the data have been analysed using BASH (BeadArray Subversion of Harshlight) tool which systematically masks beads with anomalous intensity (Cairns, Dunning et al. 2008). This tool is an adaptation of the Harshlight tool to the Illumina data to identify three major defects of arrays: localised blemishes disrupting a few probes, diffuse defects disrupting larger areas, and extended defects which invalidate the entire chip. Subsequently after evaluating the spatial artefacts,  $\beta$  values have been re-calculated as previously described.

In addition to the Beadstudio QC control, Bioconductor “Minfi” package was used as an additional QC metric (Du, Kibbe et al. 2008). Minfi analyses raw data from Beadstudio to evaluate background normalised the data and calculated M values. Beta values are heteroscedastic for highly methylated or low methylated CpG sites (Du, Zhang et al. 2010). The M-value is a Log2 ratio of the intensities of methylated probe versus unmethylated probe according to the following formula:

$$M = \text{Log2} \left( \frac{\text{Max (M intensity, 0) + } \alpha}{\text{Max (U intensity, 0) + } \alpha} \right)$$

$\text{Max}(M, 0)$  and  $\text{Max}(U, 0)$  indicate the maximum intensity value between  $M$  and 0,  $U$  and 0, respectively.  $A$  is a constant to regularise  $M$  value in order to avoid extreme changes in values that can derive from very small intensity values (between 0 and 1). Generally  $M$  values have a value between -0.5 and +0.5, which correspond to unmethylated and methylated CpG respectively.  $M$  values are widely used as metrics in downstream applications of methylation pipeline.

#### 3.3.2.2 *Subset-quantile Within Array Normalisation (SWAN) to correct InfI/InfII shift*

450K methylation array uses two different types of chemical assay on the same array, which results in different dynamic behaviour, requiring data pre-processing to make them comparable (Dedeurwaerder, Defrance et al. 2011). The normalisation method used in this study assumes a similarity of overall intensity of probes with the same number of CpGs (Figure 3.8). However distribution analysis of beta-value from different probe types often results in aberrant overall beta-value distributions. To improve the overall  $\beta$ -value distribution, raw data produced from InfI and InfII beads were normalised using SWAN (Figure 3.9). SWAN normalisation consists of a first average quantile distribution of a subset of probes with a similar number of CpGs in the probe body; the remaining probes are normalised separately by linear interpolation between the subset probes (Maksimovic, Gordon et al. 2012). Subset-quantile normalisation results not only in an improvement of the beta value distributions, but also in reduction of technical variability and better detection of differential methylation.



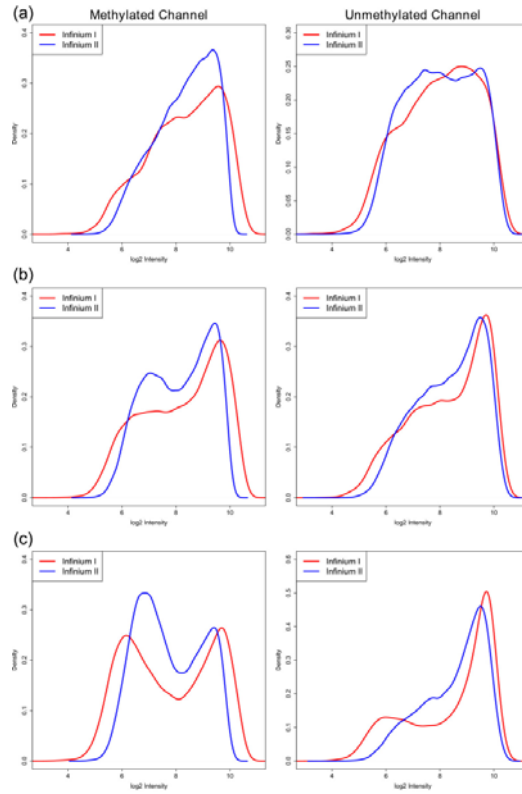


Figure 3.8 Example of similarity of intensity distributions between Infinium I and II probes with the same number of underlying CpGs. (a) One CpG in the probe body. (b) Two CpGs in the probe body. (c) Three CpGs in the probe body (Maksimovic, Gordon et al. 2012).

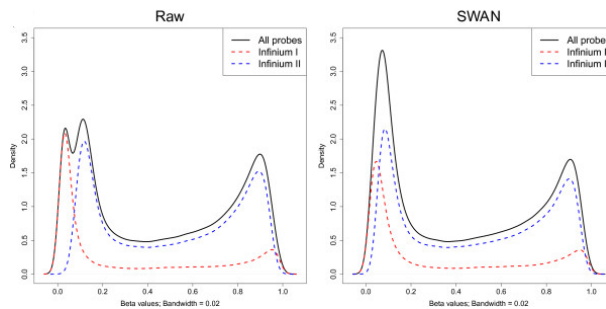


Figure 3.9 Overall beta value distribution produced by InfI and InfII results in an aberrant distribution. SWAN normalisation of the beta value distribution improves the overall distribution and reproducibility between replicates (Maksimovic, Gordon et al. 2012).

### **3.3.3 Differential methylation pattern**

The analysis of differential methylation patterns was performed using the R/Bioconductor package Limma, using moderated t-statistics (Smyth and Altman 2013). Statistical analysis on M values derived from Lumi package was performed to identify differentially genetically variant. Limma uses an empirical Bayes method to perform hypothesis tests and adjusts the p-values for multiple testing returning calculated fold changes, standard errors, t-statistics and p-values.

### **3.3.4 Detection of Copy-Number Aberrations (CNAs)**

Copy number aberrations were detected from 450K raw data as previously written (Sturm, Witt et al. 2012). Highly variant probes were excluded from the analysis; specifically probes not included between the 0.05 and 0.85 quantile of median summed values or over the 0.8 quantile of the median absolute deviation were discarded. Log-ratios and background for each sample were derived by considering the median; respectively the median value of the control and median absolute deviation of adjacent probes.

Subsequently, adjacent probes were combined to create genomic windows bigger than 100kb and smaller than 5Mb. The median probe value was calculated from raw data of each probe forming single windows. Median value were used to classify Copy-Number events for each window; values less than -0.96 indicated homozygosity, between -0.96 and -0.24 indicate homozygous deletion, between -0.24 and 0.12 a neutral event, greater than 0.72 a gain or high-level amplification.

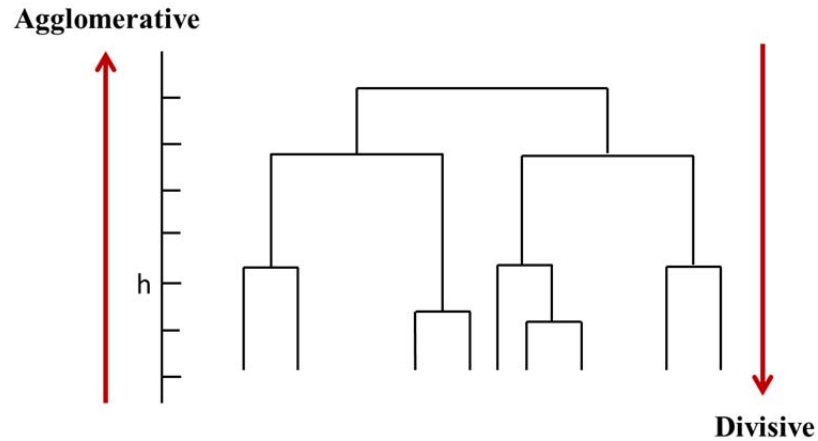
### **3.4 Unsupervised cluster analysis**

Cluster analysis is a statistical data analysis technique that explores a set of subjects aiming to create homogenous groups (cluster) such that subjects in the cluster have high similarity but differ from subjects in another cluster. The evaluation of clustering is an integral part of this study in order to organise the observed genomic data into meaningful structures.

#### **3.4.1 Hierarchical clustering**

Hierarchical clustering analysis permits a grouping of similar objects into clusters, producing a tree (called dendogram) that indicated the hierarchy of the clusters. This approach allows a visualisation of microarray and/or NGS data grouped together on the basis of similarity of their underlying features.

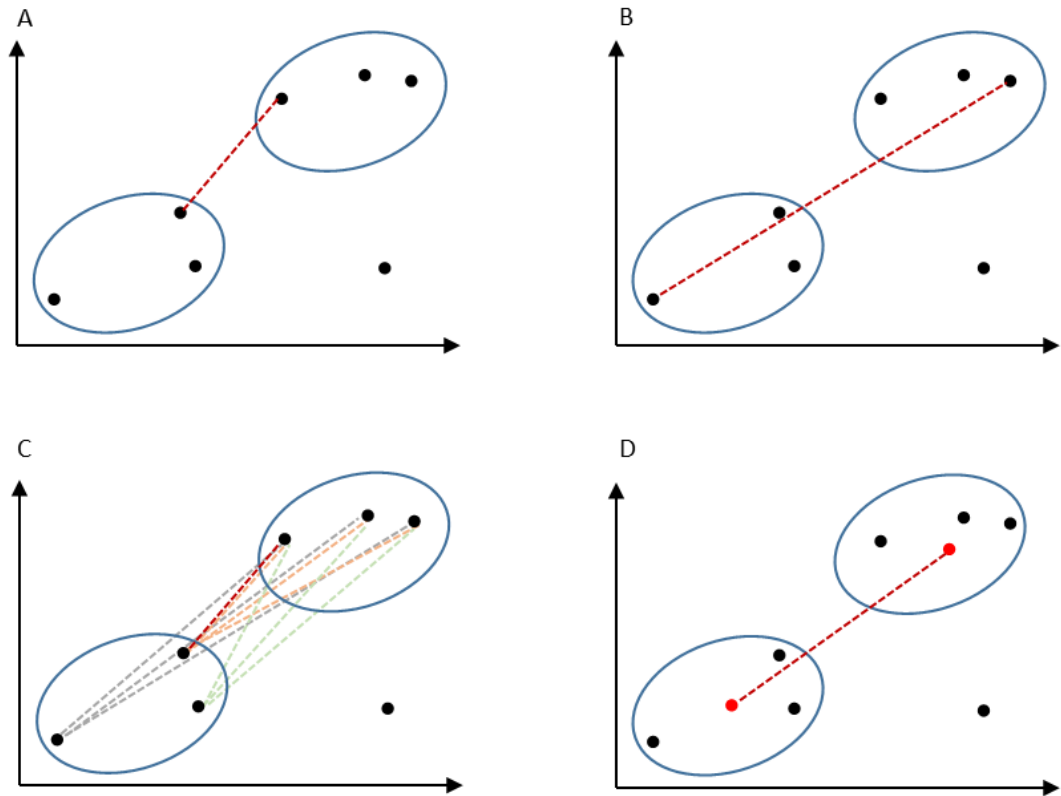
Hierarchical clustering is the result of a multistep process; Initially all objects are merged in a single cluster and gradually divided into smaller homogenous groups (divisive clustering). Cluster analysis can be also conducted in an agglomerative manner (Figure 3.10). In both cases each object when allocated to a cluster cannot be allocated to another.



*Figure 3.10 Example of hierarchical clustering. Hierarchical Clustering can be computed using agglomerative methods, in which clusters are merged together to form unique single cluster, or divisive methods, which separate objects successively into finer groupings.*

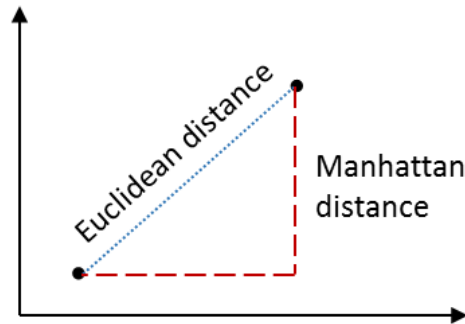
In this study agglomerative analysis has been used to explore biological differences. Distance between clusters can be calculated with four different procedures (Figure 3.11), defined by the way clusters are linked to a certain object:

1. Single linkage; the distance corresponds to the shortest distance between any two objects in the two clusters.
2. Complete linkage; this approach considers the longest distance between any two members in the two clusters.
3. Average linkage; the distance is calculated as the average distance between all paired object of the two clusters.
4. Centroid linkage; the distance equals the distance between the two geometric centres (centroids).



*Figure 3.11 Type of distance between clusters calculation methods. A) Single linkage; B) Complete linkage; C) Average linkage; D) Centroid linkage.*

In hierarchical analysis it is also important to establish dissimilarity or affinity between pairs of objects. In particular, there are two different methods: linear measurement (Euclidean) and alternative distance measurement (Manhattan). The Euclidian approach consists of a straight-linear distance; instead Manhattan is defined as sum of the absolute difference between their positions (Figure 3.12).



*Figure 3.12 Difference of methods to assess dissimilarity or affinity between pairs of objects. Euclidean function defines the distance between objects as root-sum-of-squares difference; instead the Manhattan method determines the distance as the sum of the absolute difference between their positions.*

### **3.4.2 Partitioning methods: k-means**

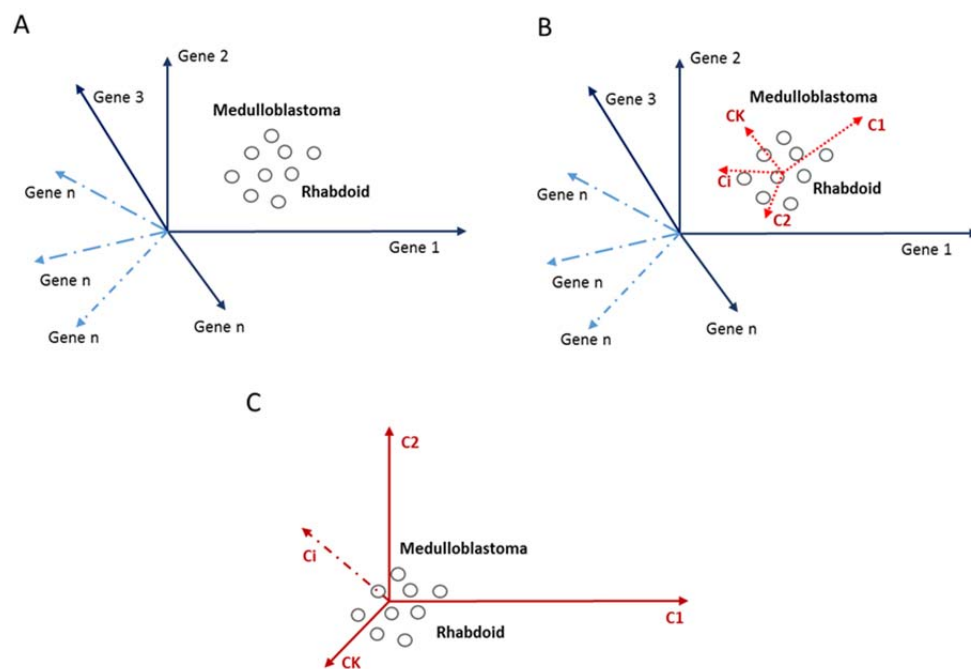
Another clustering method consists of a partitioning model, which is based on the k-means algorithm. This procedure differs from the discussed method before, since it forms homogenous clusters on the base of within-cluster variation. This vector quantisation method groups  $n$  observations from a data set in  $k$  clusters; each observation is assigned to clusters with the nearest mean.

The algorithm first randomly assigns objects to a pre-defined number of  $k$  classes and characterises the geometric centre of each cluster. After this step for each single subject Euclidean distance to the cluster centre is evaluated and each member is assigned to the nearest centroid. These steps are alternated until predetermined number of iterations is met, in this case when there is no further change in the assignment of the data points.

### **3.4.3 Principal component analysis**

Principle component analysis (PCA) is a statistical technique able to identify a limited number of key variables (principal component) of a set of variables in order to simplify

gene expression analysis. This method allows reduced dimensionality of complex data sets and reduces computational costs and error of parameter estimation. PCA is an orthogonal transformation method aimed at maximising the component vector variance of the data set. The analysis determines the direction that corresponds to maximal variance between the data points and projects the obtained data on the hyper plane orthogonal of that direction (Figure 3.13). Through this mathematical projection it is possible to obtain a simpler data set, in which the dimensions are reduced.



*Figure 3.13 Example of PCA orthogonal transformation. A) Objects from a Data set are described by a complex set of variables; B) New coordinates orthogonal to each other are set, pointing to the direction of largest variances; C) The new coordinates are used to describe the object of data set in a simpler way.*

#### 3.4.4 Non-negative matrix factorisation

Non negative matrix factorization (NMF) is a statistical analysis based on an algorithm that reduces the dimension of NGS or microarray data (Brunet, Tamayo et al. 2004, Devarajan 2008). In particular, it permits the characterisation of aggregate patterns of gene expression (metagene) from thousands of genes. In contrast to other clustering methods NMF clusters genes and samples by factorisation into matrices with non-negative entries; this results in an analysis with higher biological significance than PCA or HC.

NMF is more complex algorithmically, however it permits decomposition of data and also cross-referencing of different data set (Figure 3.14). Data sets are composed of expression levels of  $N$  genes in  $M$  samples. The data are compressed into an expression matrix  $V$  composed by values  $N \times \text{value } M$ ; rows represent expression levels of the  $N$  genes in the  $M$  samples.

Metagenes (positive linear combination of the  $N$  genes) are obtaining by factorisation of the matrix  $A$  into two matrices characterised by positive entries. Specifically, the two matrix derived are  $W$  (composed by  $N$  genes  $\times$   $k$  metagene) and  $H$  (composed by  $k$  metagene  $\times$   $M$  samples), which are approximately equivalent to the matrix  $A$ . Each  $k$  column of the first matrix defines a metagene, while each  $M$  column of  $H$  matrix corresponds to metagene expression pattern of the corresponding sample.



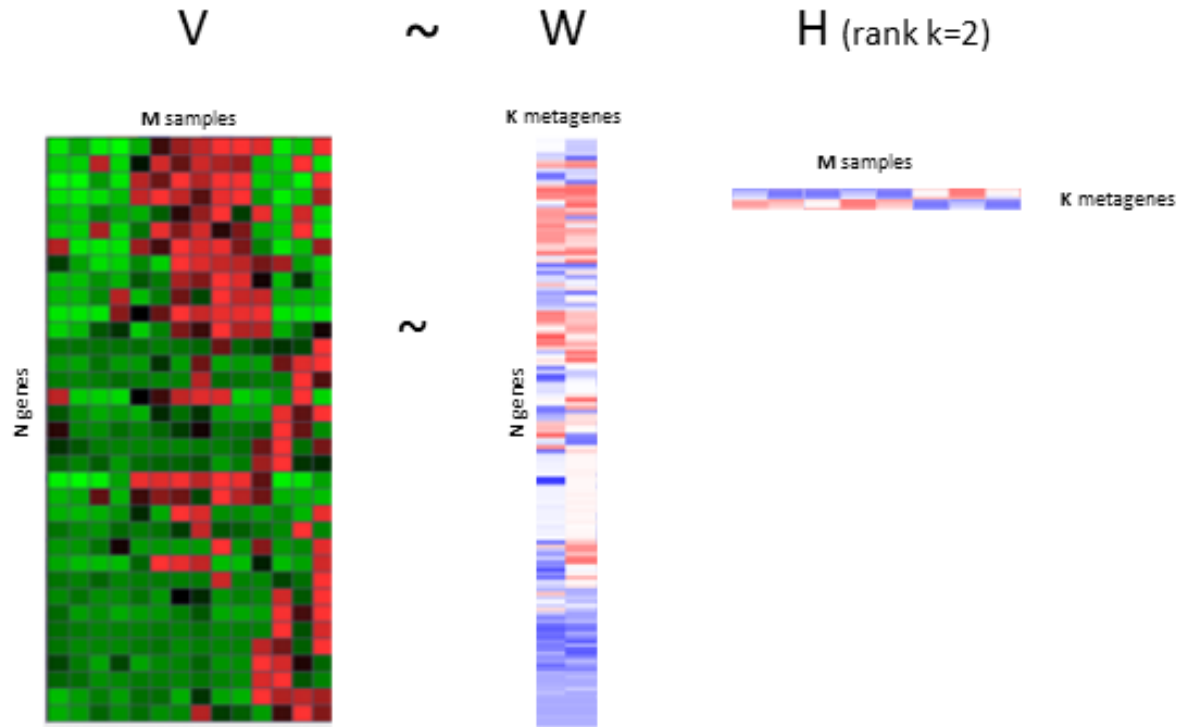


Figure 3.14 Example of NMF analysis of expression data sets composed by  $N$  genes,  $M$  samples and imposed  $k=2$ . Matrix  $V$  is a mock microarray expression data: low expressed genes are coloured in green, with the high expressed in red. Matrix  $W$  is composed of  $k$  metagenes and the number of gene and each column corresponds to the expression level of each sample for a given metagene. Matrix  $H = N$  sample with  $k$  metagenes; in this case each column represents a metagene profile for each sample. Heat colour maps are used to indicated metagene expression levels (dark blue= minimum ; dark red= maximum).

### **3.4.5 Consensus-based unsupervised clustering approach**

Unsupervised cluster analysis was undertaken in order to investigate the difference between Malignant Rhabdoid Tumours (MRT) and Medulloblastomas as well as whether there were any subgroups within the MRT cohort, defined by differences in their methylation and expression patterns. In this analysis a consensus-clustering based approach was applied Initially on RNA-seq and 450K-methylation data sets in order to determine the most robust clustering. Better definition was achieved using filtered data. The selection of optimal numbers of clusters was performed using NMF (see section 3.3.4). Each combination of 2:6 (x) metagenes and 2:6 (y) clusters were tested repeating the following steps 50 times. The most stable cluster was selected based on the average modal cluster score across all samples. The determination of the optimal number of metagenes and clusters was necessary for the consensus clustering approach.

## **3.5 Gene pathway analysis**

High-throughput sequencing techniques have greatly increased the amount of biological data, enabling comprehensive biological investigations. However, the large amount of data has to be effectively interpreted in order to accurately understand the underlying biology of the condition being studied. Modern approaches are available to simplify data sets into a smaller set, identify significant genes or proteins. In particular, in this study data sets have been interrogated to identify groups of *SMARCB1* dependent genes actively involved in the same pathway.

### **3.5.1 Gene Set Enrichment Analysis (GSEA)**

Gene Set Enrichment Analysis (GSEA) is considered a rigorous and complete statistical method to classify correlation between genes and group them into uniform classes (Subramanian, Tamayo et al. 2005). In this study GSEA analysis has been used to explore differentially expressed genes, allowing me to identify *SMARCB1* dependent

differentially regulated pathways both in primary tumours and in cell lines. GSEA analysis requires a data set pre-processing step, which ranks genes according to a phenotype or profile determining the distribution of the set of genes across *a priori* predefined gene-sets. The ranked data sets are further statistically processed and for each gene set an enrichment score (ES) is given; ES corresponds to the degree of correlation between the ranked gene list and *a priori* predefined gene-sets. Next an empirical phenotype-based permutation test is used to calculate significance levels of the observed ES, estimating p-values relative to ES null distribution. Finally, ES are normalised to the size of the original set, yielding normalised enrichment scores (NES) and false discovery rate (FDR) is computed for each NES. FDR corresponds to the estimated probability that genes identified as differentially expressed are false positive. GSEA also provides \*.txt files and visual data report, which allow a simple understanding of the enrichment of the data set of interest when compared to *a priori* predefined gene-sets.

### 3.5.2 Ingenuity pathway analysis (IPA®)

Ingenuity pathway analysis (IPA®, Ingenuity® Systems, [www.ingenuity.com](http://www.ingenuity.com)) is a web-based program which interrogates experimental data to extrapolate biological pathway information. IPA is based on the Ingenuity Knowledge Base, curated collection of biological and chemical interactions and functional annotations mainly derived from over 200,000 peer-reviewed biomedical publications or integrated from third-party databases. The database is periodically reviewed and curated by content and modelling experts. Moreover all this information has been used to generate a molecular network (interactome), a projection of physical, transcriptional and enzymatic interactions between molecules. The interactome is used to calculate networks which may simplify the understanding of the biological processes in a data set of interest, in the context of existing knowledge of interactions.

In this study RNA-seq data sets from both cells and primary tumours have been analysed using IPA in order to facilitate the molecular analysis and explore how loss of

*SMARCB1* affects pathways, focusing on those implicated in tumourigenesis. Files containing the results from the RNA-seq experiment were uploaded into the Ingenuity Pathways Knowledge Base. After having applied a cut-off of 2-fold, the molecules retrieved were compared to the IPA global molecular network based on their connectivity, using Fischer Exact tests and a p value of  $<0.05$ . IPA analysis produced a score for each interaction network produced, based on the number of eligible molecules present in the network, the total number of eligible molecules investigated and the total number of molecules present in the Ingenuity Knowledge Base. Finally, the networks recovered by the analysis have been ranked based on the IPA score and further analysed.

### **3.5.3 Discovery Database for Annotation, Visualisation and Integrated (DAVID)**

The Database for Annotation, Visualisation and Integrated Discovery (DAVID) is a web base tool which strategically extracts biological meanings from a gene data set (Dennis, Sherman et al. 2003). DAVID is based on DAVID knowledgebase comprehensive of 40 heterogeneous gene annotation resources (NCBI and UniProt annotation).

A selected list of genes based upon RNA-seq data from primary material or cell lines are submitted to the online DAVID tool. These are first scored based upon gene-gene similarity with the *a priori* assumption that genes are functionally related. Next the data sets are functionally grouped into biological modules, corresponding to biological processes. In biological systems genes can be active in more than one processes, therefore DAVID's agglomerate algorithm allows the allocation of a gene to more than one functional group. Statistically significant results are determined by applying Fisher's exact test ( $P \leq 0.05$ ) which is adjusted for multiple testing.

## **Chapter 4      Establishing the Role of *SMARCB1* Loss in Malignant Rhabdoid Tumour Cells**

## 4.1 Introduction

Rhabdoid Tumours are characterised by loss of *SMARCB1* gene, they do not demonstrate any other consistent mutation or chromosomal aberrations (Hasselblatt, Isken et al. 2013). The discovery of *SMARCB1* as the single cause of tumourigenesis in Rhabdoid Tumours has highlighted this gene as a potent tumour suppressor. Research into the effects of re-expression of *SMARCB1* in a Rhabdoid cell line has been previously attempted; yet little is known about its consequence on the whole genome. Previous work has demonstrated that *SMARCB1* re-expression causes cell cycle arrest in G1, following upregulation of genes such as *CDKN2A* and *CDKN1A* and recruitment of *Rb* (Oruetxebarria, VenturSMARCB1 et al. 2004, Chai, Charboneau et al. 2005, Kuwahara, Charboneau et al. 2010). It has also been demonstrated that *SMARCB1* plays an important role in chromatin remodelling and its re-expression increases the binding to *CDKN2A* and *CDKN1A* promoters by changing the methylation status of histone marks H3k4 and H3k27.

It has been speculated that *SMARCB1* loss not only affects gene expression but also methylation. In fact, *SMARCB1* is part of the SWI/SNF complex and several SWI/SNF-related factors have been associated with transcriptional silencing through DNA methylation (Dennis, Fan et al. 2001). Specifically it has been hypothesised that a selective interaction of these complexes with MeCP2 might change the chromatin structure to a repressive state (Chandler, Guschin et al. 1999). Moreover, in *SMARCA4* and *BRM* deficient carcinoma cells *CD44* and *E-cadherin* are hyper-methylated and induction of the two ATPase domains, as well as 5-aza-2-deoxycytidine treatment, re-establishes the methylation and expression level (Banine, Bartlett et al. 2005). In Rhabdoid cells, the implication of *SMARCB1* in the methylation machinery has been studied in only a few cell lines, in which re-expression of *SMARCB1* has been combined with the 5-aza-2'-deoxycytidine demethylating agent (Muhlish, Schwering et al. 2005, Kia, Gorski et al. 2008). These two isolated experiments showed how re-establishment of *SMARCB1* function affected the methylation status and expression of *CDKN2B* and *CDKN2A* in MON cells and *RASSF1A* in G401 and BT-16 cells. Notably, *SMARCB1* re-expression directly evicts the prebound PRC1 and PRC2 component of the PcG

complex and subsequently causes loss of methylation of H3k27 at the *CDKN2A* promoter locus (Kia, Gorski et al. 2008). 5-aza-2-deoxycytidine treatment in the renal Rhabdoid G401 cell line results in hypomethylation and re-expression of *RASSF1A* (Muhlish, Schwering et al. 2005), a gene found often to be hypomethylated in a variety of cancer (Dammann, Schagdarsurengin et al. 2003). In contrast, *RASSF1A* expression in the Atypical Teratoid Rhabdoid Tumours derived BT16 cell line does not alter upon 5-aza-2-deoxycytidine treatment, suggesting a different methylation status dependent on the location.

Other subunits of the SWI/SNF complex have been found mutated in a variety of cancer. However, the most frequently mutated subunit is not *SMARCB1* but the ATPase subunit *SMARCA4*, suggesting its role as potent tumour suppressor. In *SMARCA4* deficient tumour cells re-expression of wild-type *SMARCA4-SMARCA4* causes induction of *CD44* and of *CDKN2A* and cell cycle arrest by RB recruitment (Strobeck, Knudsen et al. 2000, Strobeck, DeCristofaro et al. 2001). Moreover heterozygous inactivation of *SMARCA4-SMARCA4* in mice predisposes to ependymal and apocrine tumours development (Bultman, Gebuhr et al. 2000). All these effects have been also observed when *SMARCB1* is restored in Rhabdoid cell lines, suggesting *SMARCB1* and *SMARCA4* subunits might act together in maintenance of cell homeostasis. Recently, a germline nonsense mutation and somatic inactivation of *SMARCA4/SMARCA4* in *SMARCB1* positive Rhabdoid patients has been also observed, supporting the hypothesis of a Rhabdoid Tumour Predisposition Syndrome (Hasselblatt, Gesk et al. 2011). The presence of this mutually exclusive mutation suggests that in Rhabdoid cells tumorigenesis might be the result of disruption of remodelling activities of these two SWI/SNF subunits.

This part of the study was focused upon building Rhabdoid cell models in which *SMARCB1* was stably expressed. Moreover, aberrant DNA methylation has been also explored, by treatment of cell lines with 5-aza-2'-deoxycytidine specific inhibitors of DNA methylation. The role of *SMARCA4* in Rhabdoid development has also been investigated, blocking *SMARCA4* activity by insertion of a dominant negative *SMARCA4*. In order to have a wider and exhaustive catalogue of malignant Rhabdoid

events, four cell lines were subjected to experiments: G401, STA-WT1 A204, and CHLA-266 derived respectively from primary Rhabdoid tumour of the kidney, liver and the brain. These four cell lines represent the most common location of tumours, giving the possibility to study the role of *SMARCB1* at different location in tumour development. Furthermore G401, A204 and STA-WT1 cell lines have previously been established as a model to study *SMARCB1* re-expression; however this approach has never been attempted in cells derived from Atypical Teratoid Rhabdoid primary tumour.



## 4.2 Aims

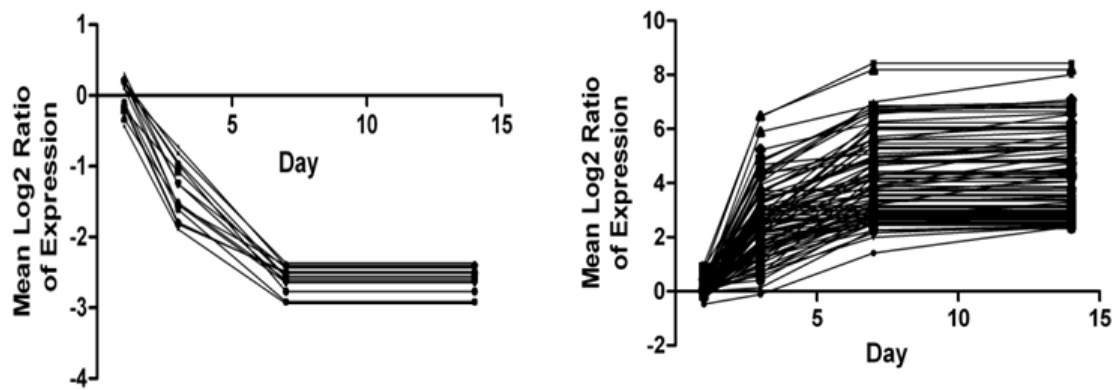
The study reported in this chapter aimed to:

- Assess the effect of *SMARCB1* re-expression in gene expression, cell morphology, and cell growth and at the chromatin level in Rhabdoid cells derived from different location.
- Investigate the methylation status in Rhabdoid cells and how it changes following treatment with 5-aza-2'-deoxycytidine.
- Understanding the role of *SMARCA4* in Rhabdoid Tumours.
- Develop and validate functional genomic Rhabdoid cell line models.

### **4.3 Re-expression of *SMARCB1* in established Malignant Rhabdoid cells lines**

#### **4.3.1 *SMARCB1* causes growth arrest and senescence cell formation in Rhabdoid cell lines**

In order to investigate the tumourigenic mechanism involved in Rhabdoid tumour, we first established a stable *SMARCB1* expression system in Rhabdoid cells via a lentiviral approach. Previous work reported early effects (usually 72h) of *SMARCB1* re-expression, however late ones have not yet characterised. In this study, we performed a preliminary time-course induction of *SMARCB1* in the G401 cell line followed by microarray analysis, in order to describe the late responses and the non-transient change in cells. We observed that *SMARCB1* re-expression caused a general disruption of expression up until the 7<sup>th</sup> day post infection, with particular emphasis on expression patterns involving cell cycle control and differentiation (Figure 4.1). In fact, microarray analysis of the RNAs isolated at various times showed that hundreds of genes were both up and down regulated, but generally that expression was stable after the 7<sup>th</sup> day following *SMARCB1* re-expression. Given this result RNA from all four Rhabdoid cell lines was collected at the 7<sup>th</sup> day post infection.



*Figure 4.1 SMARCB1 induction affects gene expression in the G401 cell line (time course). Microarray analysis evidenced hundreds of differentially expressed genes, here shows as a log2 ratio of expression compared to the empty vector. A total of 1256 genes were down-regulated (A) and 374 were up regulated (B) following SMARCB1 reexpression.*

Following infection, the shape of *SMARCB1* induced cells appeared different from vector control infectants by visual inspection. Specifically, the cells acquired flat and large morphology, with an increased adhesion to the extracellular matrix. This cellular re-arrangement in Rhabdoid cells has been often associated with cellular senescence (Versteeg, Medjkane et al. 2002, Reincke, Rosson et al. 2003, Oruetxebarria, Venturini et al. 2004). *SMARCB1* re-expression not only resulted in flattened appearance and senescence-associated large shape (Figure 4.2) but also in the dramatic decrease of cell proliferation, this effect was observed from the 3<sup>rd</sup> day post infection onwards (Figure 4.3).

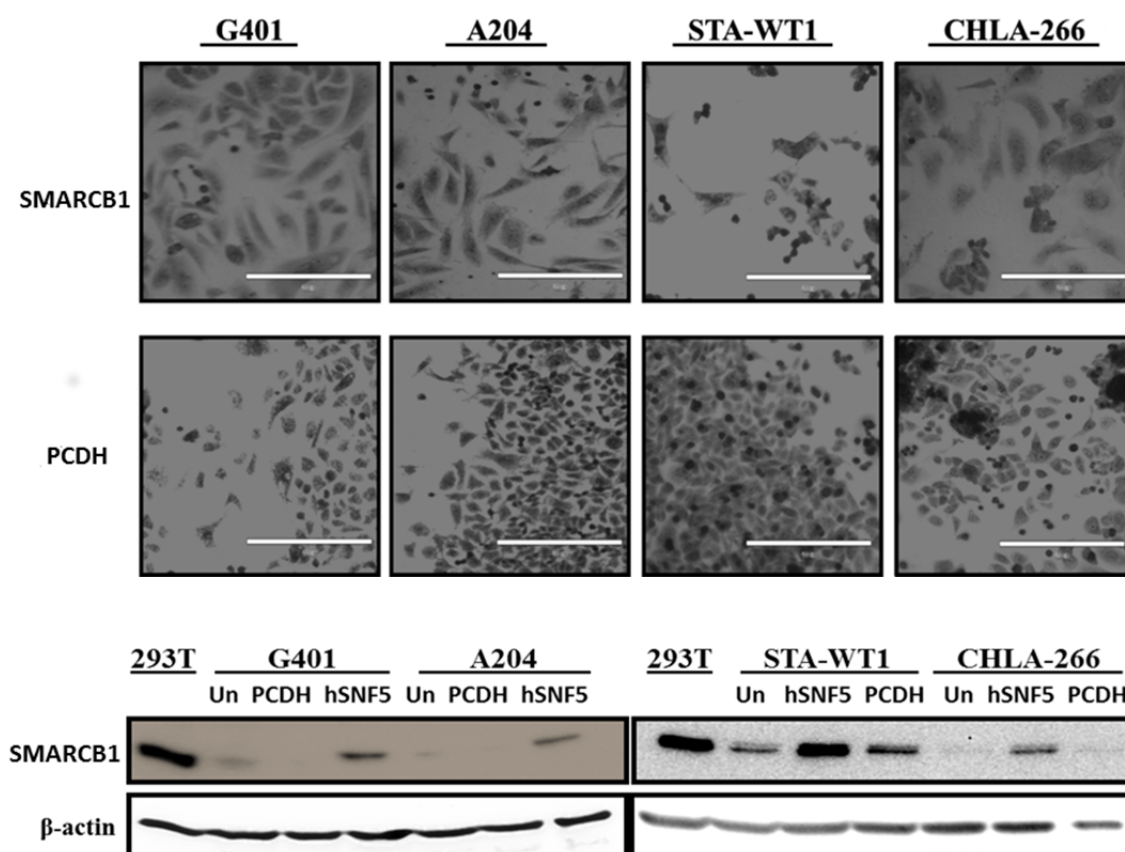
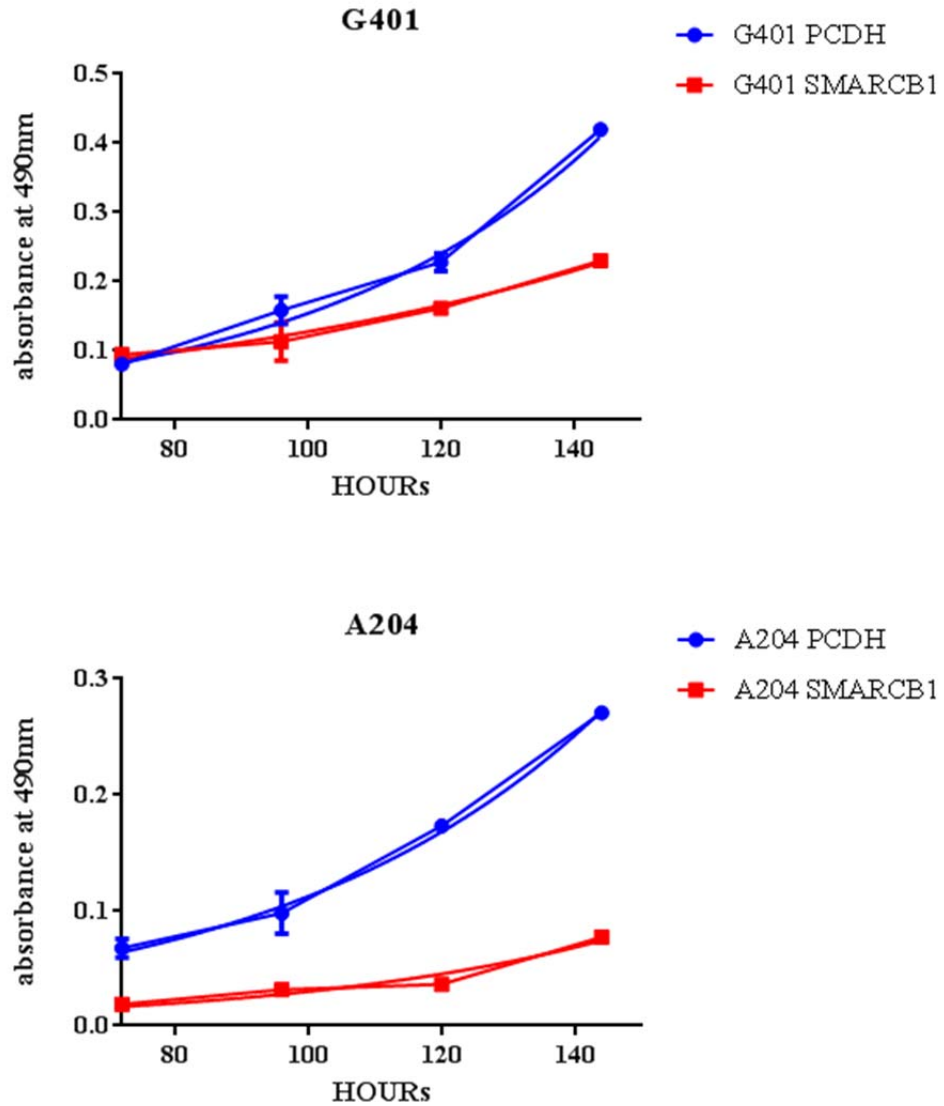
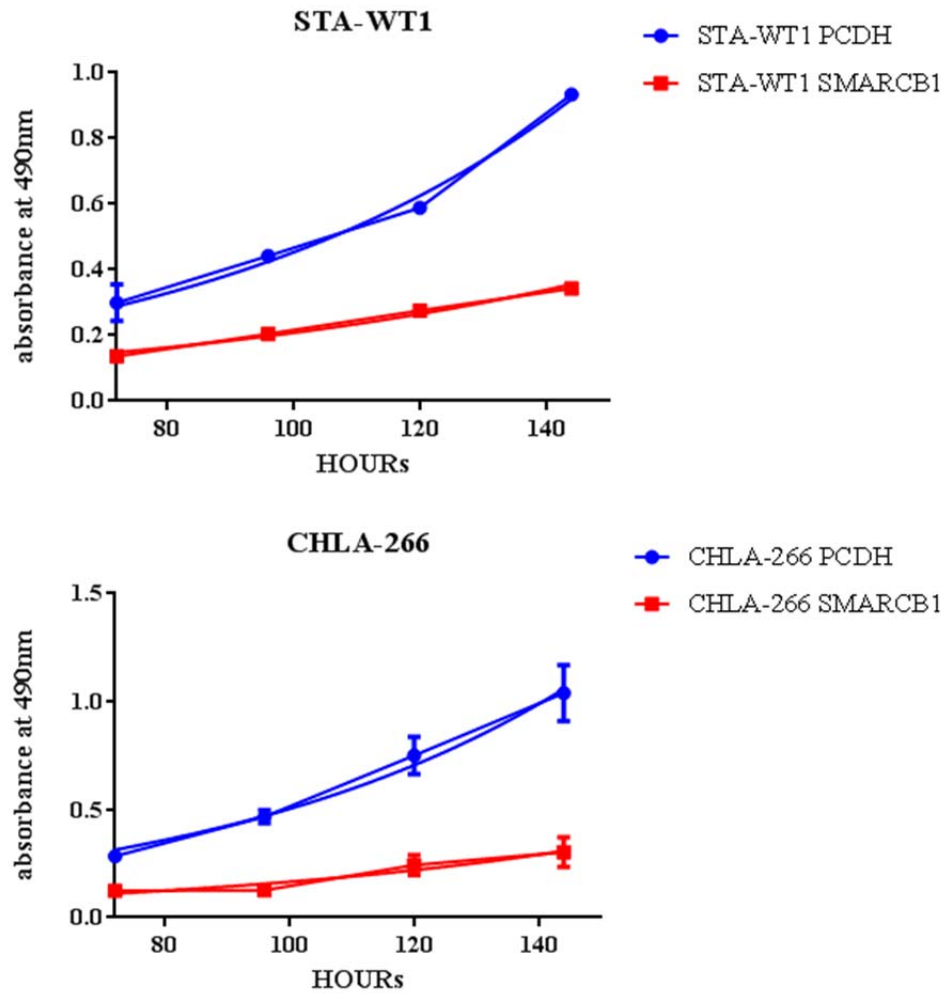


Figure 4.2 Change in cell morphology after SMARCB1 induction in Rhabdoid cell line (7 days). After re-expression of SMARCB1 the cell morphology has visibly changed: the cells are bigger, flat with evident changes in shape. Western blot from cell lysates obtained from uninfected cell (Un), from cell infected with *pcdh-mcs1-ef1-puro* (PCDH) and SMARCB1-*pcdh-mcs1-ef1-puro* (hSNF5). Proteins were extracted at 7 days post infection. Protein from the 293t cell line has been used as positive control. STA-WT1 cell line is characterized by SMARCB1 point mutation and retains SMARCB1 protein.



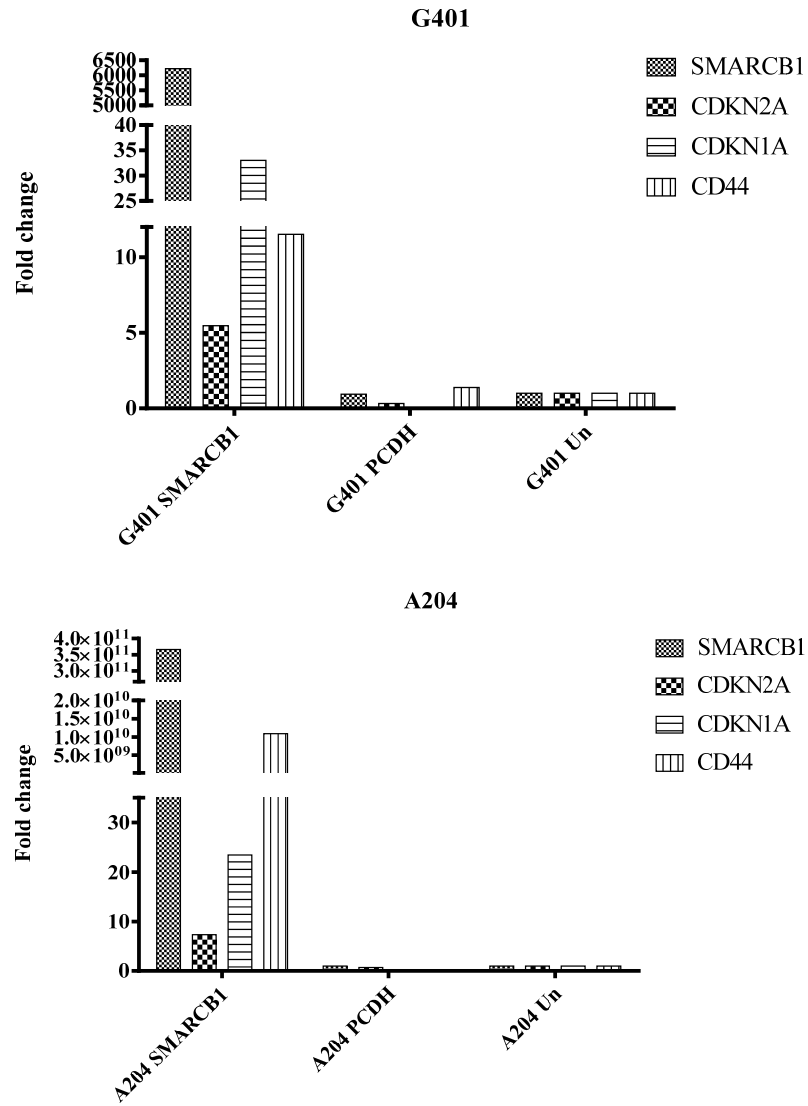
*Figure 4.3 Effect of SMARCB1 re-expression on survival (growth) in G401 and A204 Rhabdoid cells. SMARCB1 re-expression causes significant growth arrest within 96 hours post infection, compared to the control-vector infectants. The Rhabdoid cells were infected with pcdh-mcs1-efl-puro (PCDH) and SMARCB1-pcdh-mcs1-efl-puro (SMARCB1) and measurements taken between 72h-144 hours post-infection. The graphs show results of three individual experiments.*



*Figure 4.4 Effect of SMARCB1 re-expression on survival (growth) in STA-WT1 and CHLA-266 Rhabdoid cells. SMARCB1 re-expression causes significant growth arrest within 96 hours post infection, compared to the control-vector infectants. The Rhabdoid cells were infected with pcdh-mcs1-efl-puro (PCDH) and SMARCB1-pcdh-mcs1-efl-puro (SMARCB1) and measurements were taken post infection (72h-144 hours). In the graphs are indicated results of three individual experiments.*

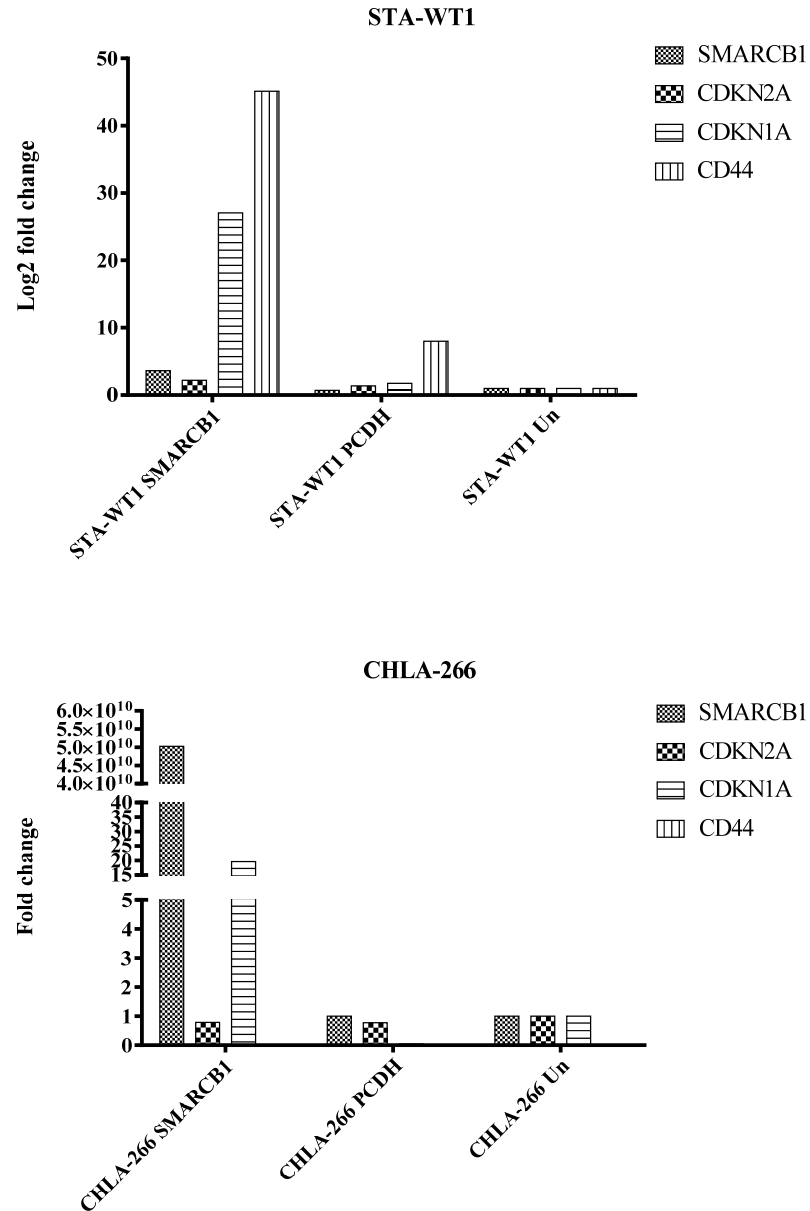
#### **4.3.2 *SMARCB1* induces *CDKN2A*, *CDKN1A* and *CD44* over-expression in Malignant Rhabdoid Tumour cell lines**

*SMARCB1* re-expression in Malignant Rhabdoid Tumours deregulates expression of genes that control cell cycle regulation such as *CDKN2A*, *CDKN1A* and *CD44* (Kuwahara et al, 2010; Kuwahara et al, 2013; unpublished data). The RNA levels of these gene were analysed 7 days post infection and were measured by quantitative RT-PCR. The cell lines G401 and A204 cell line show an increase of *CDKN2A* (5.4 fold and 7.3 fold respectively), *CDKN1A* (33 fold and 23.4 fold respectively) and *CD44* (11.1 fold and  $1 \times 10^6$  fold respectively) compared to the vector control infectants and untreated cells (Figure 4.5). The expression values obtained are similar to those previously observed in our time course (data not indicated) and those in the literature (Kuwahara, Charboneau et al. 2010). STA-WT-1 showed a significant increase of *CDKN1A* (27 fold) and *CD44* expression (45.1 fold) similar to G401 and A204 but no substantial increase in *CDKN2A* expression (2.2 fold). CHLA 266 showed increases in expression of *CDKN1A* expression similar to the other cell lines (19.5 fold), but no increase in *CDKN1A* and expression of *CD44* was undetectable in both untreated cells and the empty vector control (Figure 4.6).



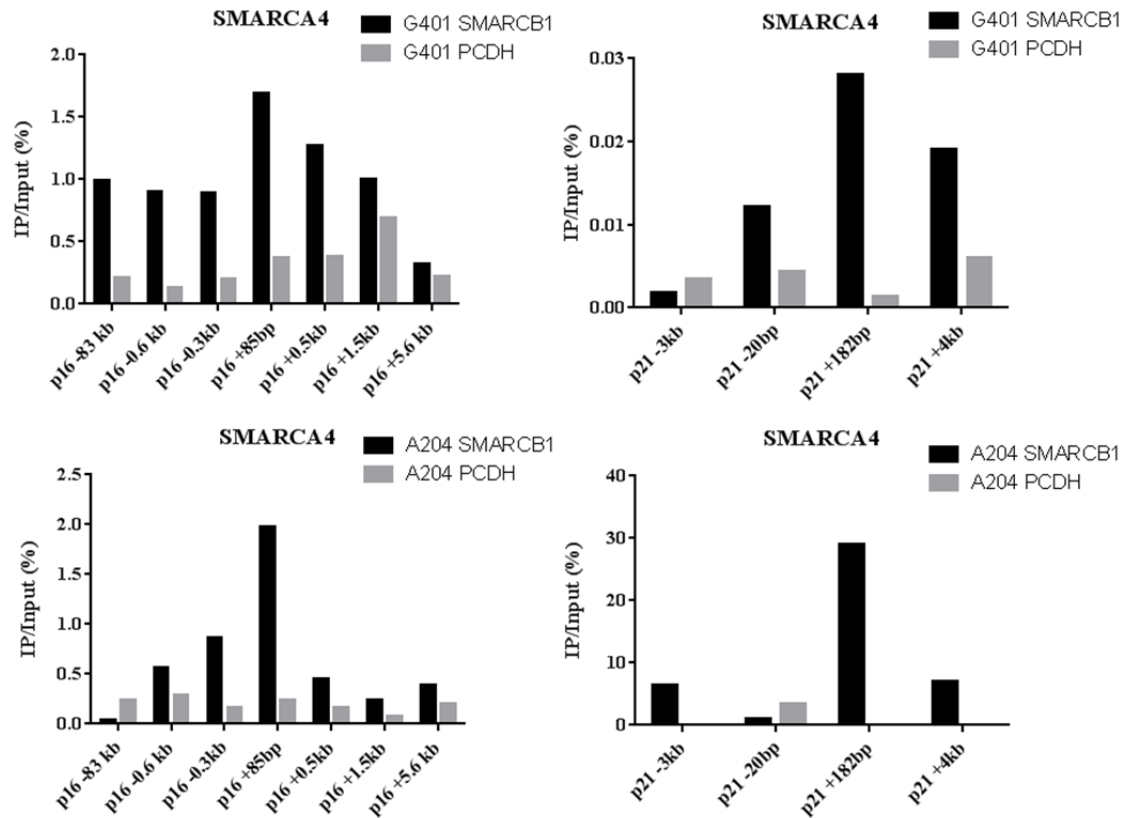
*Figure 4.5 SMARCB1 re-expression induces changes in gene expression in G401 and A204. RNA was extracted at 7 days post infection with pcdh-mcs1-ef1-puro (vector control) and SMARCB1-pcdh-mcs1-ef1-puro. The mRNA levels were measured by real-time qRT-PCR and normalised to beta-2-microglobulin. Values are the mean of 3 independent measurements. Untreated cells (Un) were used as a control.*



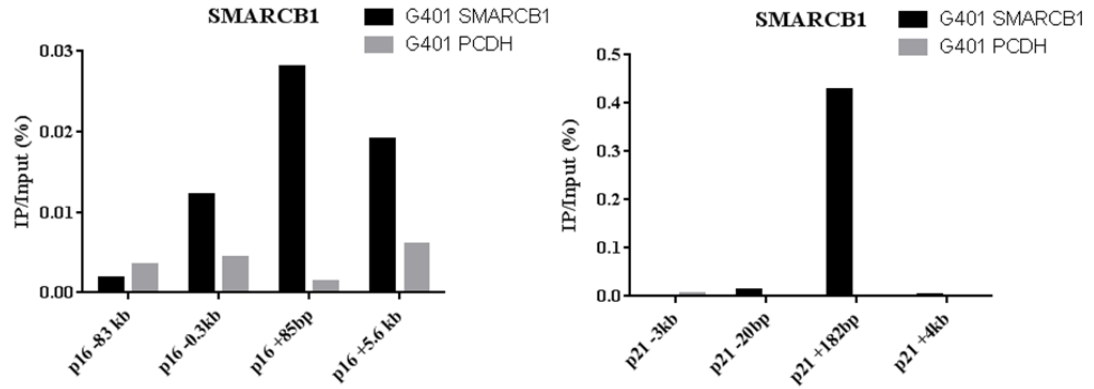


*Figure 4.6 SMARCB1 re-expression induces changes in gene expression in G401 and A204. RNA was extracted at 7 days post infection with pcdh-mcs1-ef1-puro (vector control) and SMARCB1-pcdh-mcs1-ef1-puro. The mRNA levels were measured by real-time qPCR and normalised for beta-2-microglobulin. Values are the mean of 3 independent measurements. Untreated cells (Un) have been used as control.*

Changes in the recruitment of the SWI/SNF complex caused by *SMARCB1* re-expression in Malignant Rhabdoid Tumours were also investigated at the *CDKN2A* and *CDKN1A* locus. Chromatin immunoprecipitation was performed on immunoprecipitated DNA using antibodies for SMARCB1, SMARCA4, H3k4me and H3k27me and appropriate control IgGs at 7 days post infection in the A204 and G401 cell line. Re-expression of *SMARCB1* alters binding at a locus within 1 kb of *CDKN2A* and *CDKN1A* of the TSS (Transcriptional Start Site) with maximal enrichment at the TSS in both A204 and G401 cell lines (Figure 4.7). Similar results were obtained using a SMARCB1 antibody, in contrast with results previously described by Kuwahara (Kuwahara, Mora-Blanco et al. 2013), who reported a general increase throughout the whole region (Figure 4.8). In our experiment *SMARCB1* re-expression increases the levels at the *CDKN2A* and *CDKN1A* TSS of H3K4me3 (Figure 4.10), a chromatin mark associated with gene activation, whilst inhibiting the repressive H3K27me3 marks. These results indicate that SMARCB1 actively binds to the TSS in combination with the SMARCA4 ATPase subunit, promoting methylation of the H3K4 in a stable manner (Figure 4.9).



*Figure 4.7 SMARCB1-dependent recruitment of SMARCA4 at the CDKN2A (p16) and CDKN1A (p21) locus after SMARCB1 re-expression in G401 and A204 cells. At 7 days after SMARCB1 re-expression. Chromatin was extracted for Chip assay, performed using antibody against SMARCA4 and precipitated DNA was measured by Q-PCR at various locations spanning the CDKN1A and CDKN2A locus; Values are the mean of triplicates. Empty vector infected cells (pCDH) were used as a control.*



*Figure 4.8 Recruitment of SMARCB1 at the CDKN2A (p16) and CDKN2A (p21) locus after SMARCB1 re-expression in G401 cells. At 7 days following SMARCB1 re-expression. Chromatin was extracted for Chip assay, performed using antibody against SMARCB1 and precipitated DNA was measured by Q-PCR at various locations spanning the CDKN1A and CDKN2A locus; Values are the mean of triplicates. Empty vector infected cells (pCDH) were used as a control.*

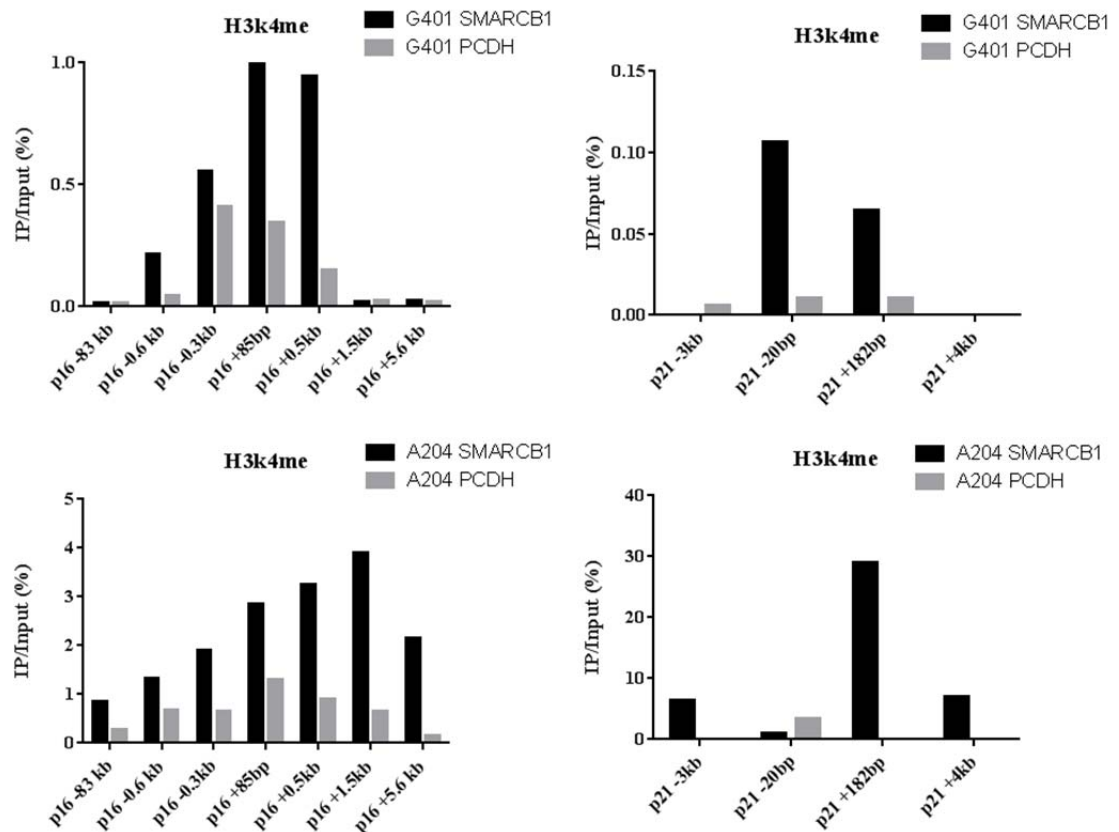


Figure 4.9 Recruitment of H3k4me3 histone modification at the CDKN2A (p16) and CDKN2A locus after SMARCB1 re-expression in G401 and A204 cells. At 7 days following SMARCB1 re-expression. Chromatin was extracted for Chip assay, performed using antibody against H3k4me3 and precipitated DNA was measured by Q-PCR at various locations spanning the CDKN1A and CDKN2A locus; Values are the mean of triplicates. Empty vector infected cells (pCDH) were used as a control.

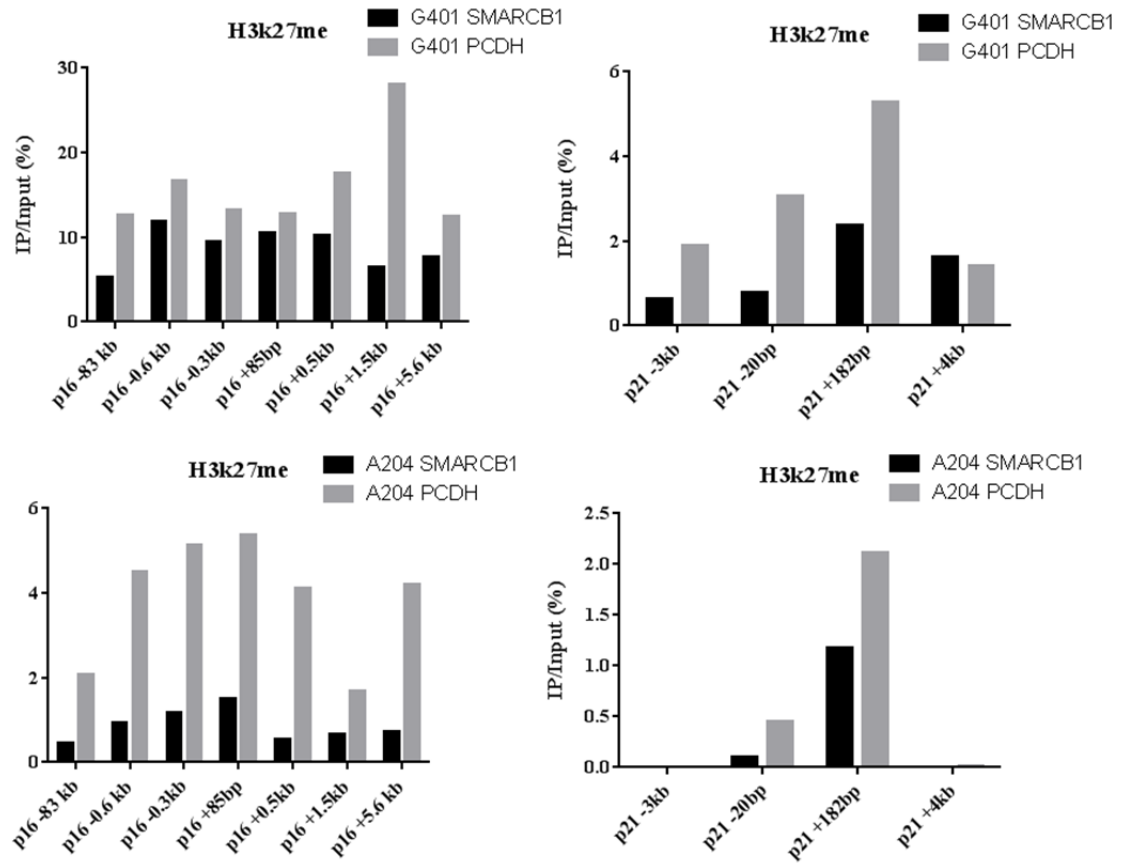


Figure 4.10 Recruitment of H3k27me3 histone modification at the CDKN2A (p21) and CDKN2A locus following SMARCB1 re-expression in G401 and A204 cells. At 7 days following SMARCB1 re-expression Chromatin was extracted for Chip assay, performed using antibody against H3k27me and precipitated DNA was measured by Q-PCR at various locations spanning the CDKN1A and CDKN2A locus; Values are the mean of triplicates. Empty vector infected cells (pCDH) were used as a control.

#### **4.4 Effects of 5-aza-2-deoxycytidine treatment on methylation status of Malignant Rhabdoid cells**

In this part of the study all four Malignant Rhabdoid Tumour (MRT) cell lines (G401, A204, CHLA266, STA-WT1) have been treated with the DNA demethylating agent 5'-aza-2'-deoxycytidine (5-azaCdR), in order to analyse to what extent the gene expression profile of MRT is controlled by methylation status and to investigate if changes in gene expression upon SMARCB1 re-expression are the result of methylation status modification. 5-azaCdR is a demethylating agent that irreversibly binds DNA methyltransferase enzymes (DNMTs), resulting in a block of their methyltransferase function. As a result, 5-azaCdR induces significant demethylation in human cell lines (Jones 2012).

We investigated the effect in methylation in three genes: *RASSF1A*, *DAZL* and *HTATIP2* were examined; the first gene has been found up-regulated in G401 cell line in response of 5-azaCdR (Muhlisch, Schwering et al. 2005). The latter in Medulloblastoma cell lines (Anderton, Lindsey et al. 2008). Up-regulation of *DAZL* and *HTATIP2* was observed in all cell lines upon treatment with 5-azaCdR. *RASSF1A* was modestly upregulated in A204, STA-WT1 and CHLA-266 but not in G401 (Figure 4.11).

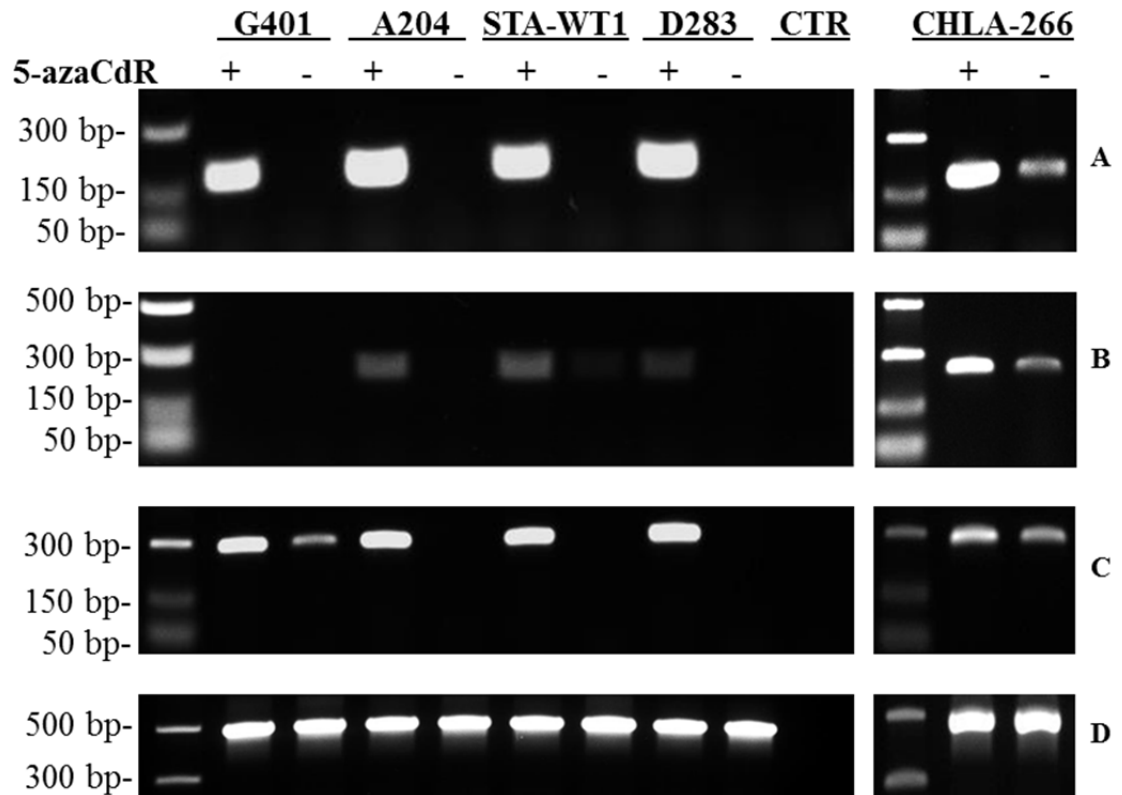
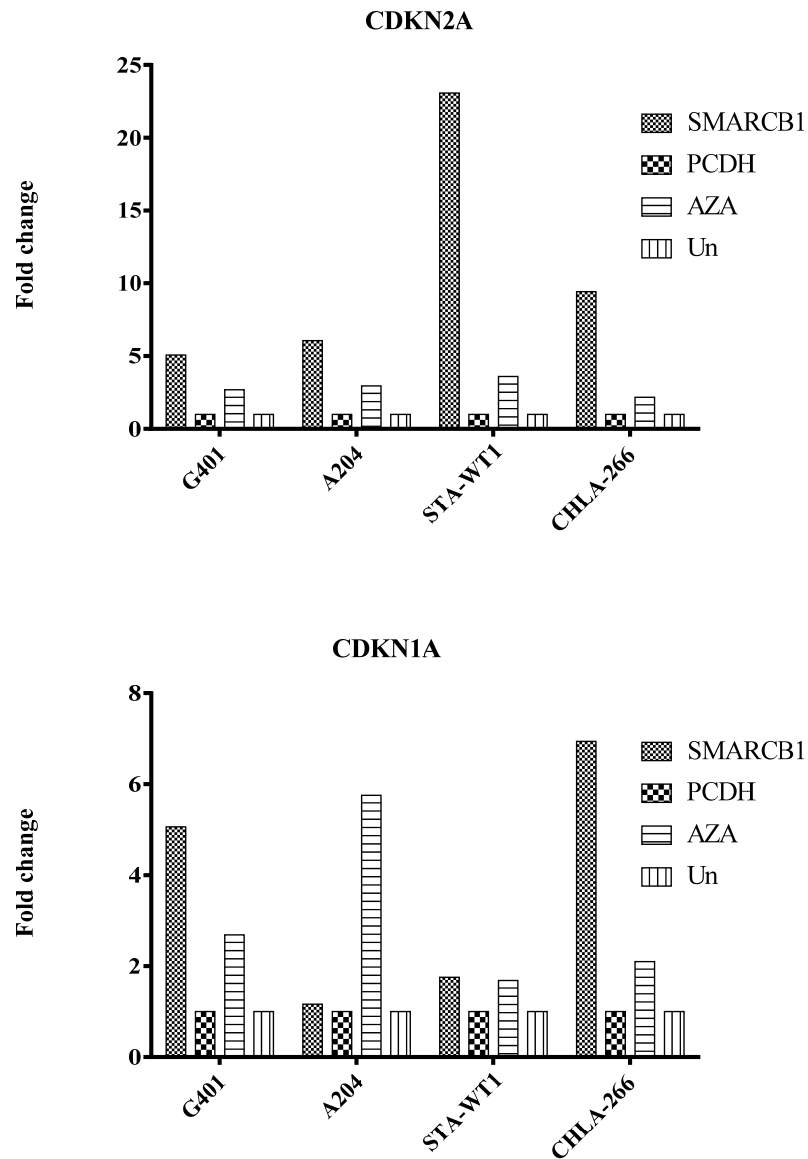


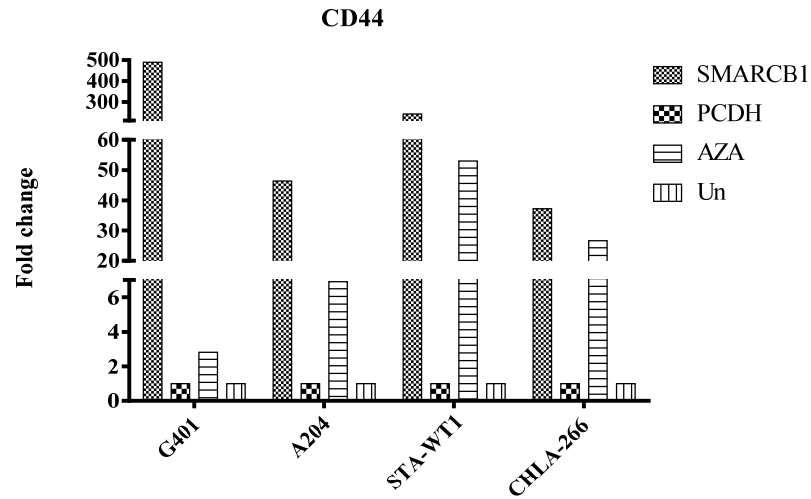
Figure 4.11 RT-PCR for *DAZL* (A), *RASSF1A* (B), *HTATIP2*(C) and *GAPDH* (D) in Malignant Rhabdoid Tumour cells treated with 5-azaCdR. D283cell (Medulloblastoma) line was used as control, as well as a no template. 5-azaCdR treatment causes upregulation of these three genes through a methylation dependent process.

We also analysed the effect of 5-azaCdR treatment on the expression of *SMARCB1* dependent genes *CDKN2A*, *CDKN1A* and *CD44* in order to understand the role of methylation in *SMARCB1* dependent regulation (Figure 4.12 and Figure 4.13). Interestingly, comparative analysis of the mRNA level between *SMARCB1* re-expressing infected Malignant Rhabdoid Tumour cells and 5-azaCdR treated cells revealed equivalent changes in expression, suggesting *SMARCB1* expression may induce demethylation and therefore changes in expression





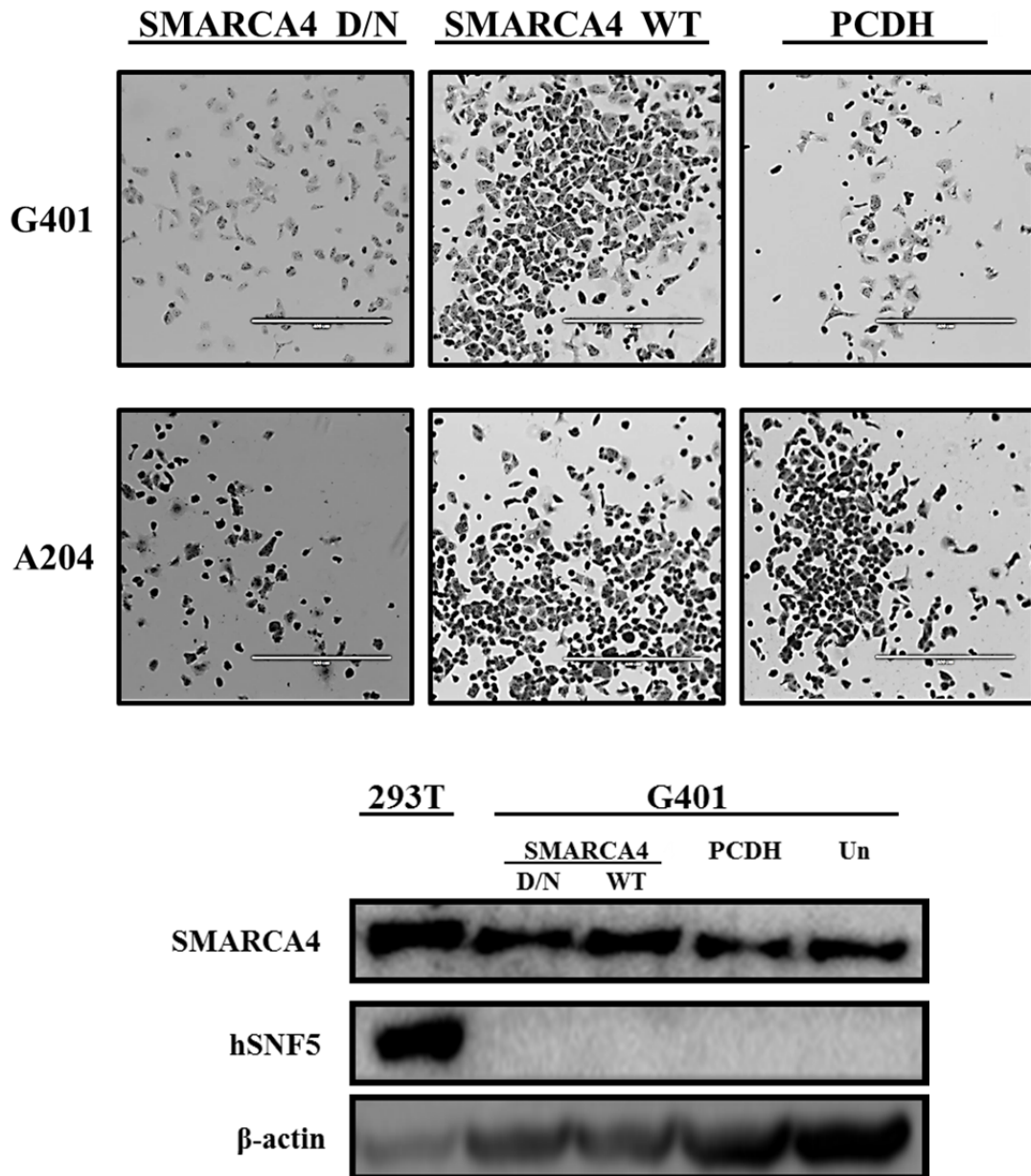
*Figure 4.12 5-azaCdR induces changes in CDKN2A and CDKN1A expression in Rhabdoid cells in a similar manner to that observed following SMARCB1 re-expression. RNA was extracted at 5 day after treatment. The mRNA levels were measured by real-time qPCR and normalised to beta-2-microglobulin. Values are the mean of 3 independent measurements. Untreated cells (Un) and empty vector infected cell (PCDH) have been used as controls expression is quoted relative to the untreated control.*



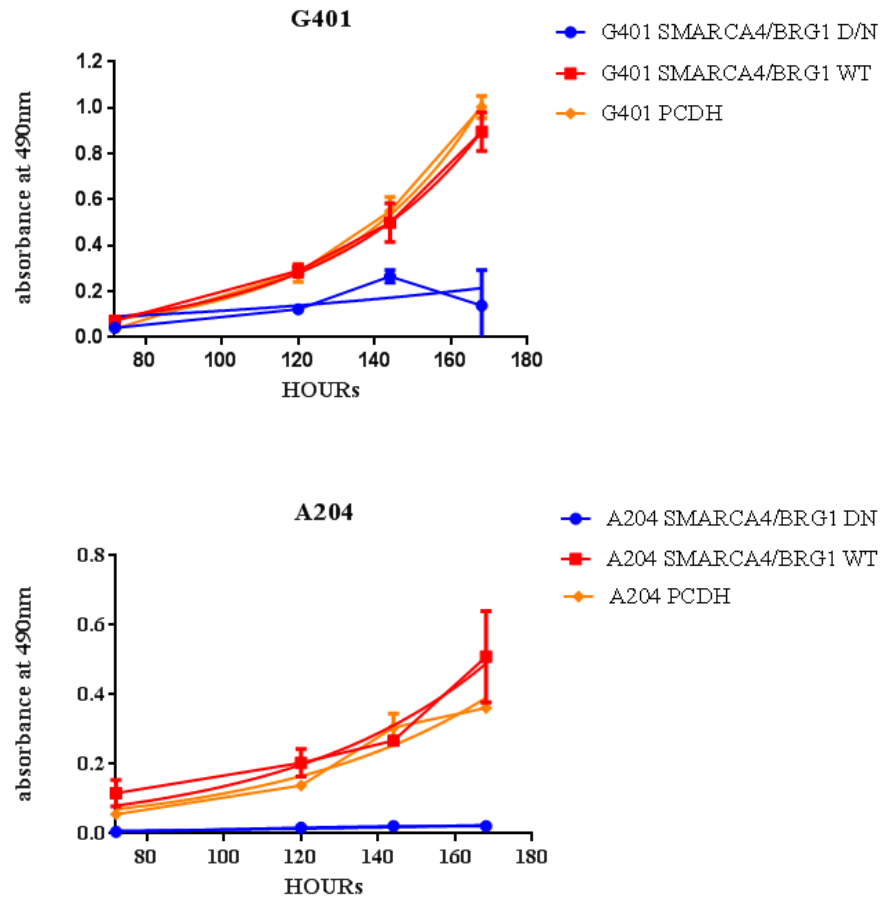
*Figure 4.13 5-azaCdR induces change in CD44 expression gene expression in Rhabdoid cells in a similar manner to that observed in SMARCB1 re-expression. RNA was extracted at 5 day after treatment. The mRNA levels were measured by real-time qPCR and normalised to beta-2-microglobulin. Values are the mean of 3 independent measurements. Untreated cells (Un) and empty vector infected cell (PCDH) have been used as controls expression is quoted relative to the untreated control.*

#### **4.5 SMARCA4 ATPse subunits acts as proto-oncogene in Malignant Rhabdoid Tumour**

The role of the SMARCA4 subunit in Rhabdoid Tumours development has been also investigated. A dominant negative *SMARCA4* (K785R) was re-expressed in G401 and A204 cell lines using lentiviral infection. We observed after 36 hours post infection a change in the cellular shape of the *SMARCA4* DN infectants: the cells appeared smaller with a dense cytoplasm while the organelles evolved into a tight and packed conformation (Figure 4.14). These features are more related with apoptosis than senescence (Elmore 2007). Also we observed a drastic decrease of cell growth following *SMARCA4* inactivation (Figure 4.15).

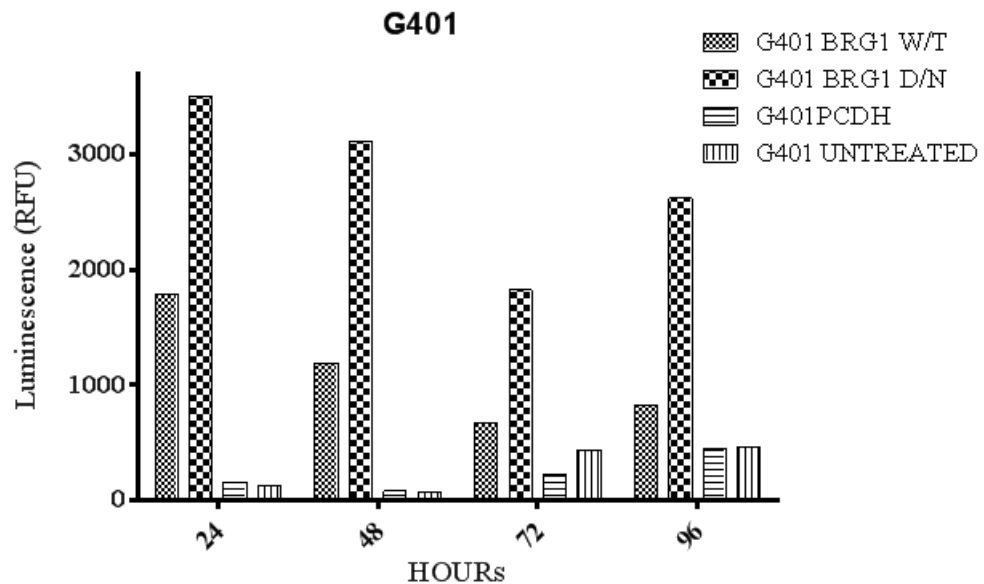


*Figure 4.14 Change in cell morphology after SMARCA4 induction in Rhabdoid cell line (48 hours post infection). The cell morphology has visibly changed: the cells are visibly smaller in size and the cytoplasm is dense in comparison to the SMARCA4 W/T and the empty vector control infectants. SMARCA4 D/N protein validation by western blot was possible just in G401, due to the poor quality of the protein extracted. Protein from 293t cell line has been used as positive control. Untreated cells (Un) and empty vector infected cell (PCDH) have been used as negative control.*

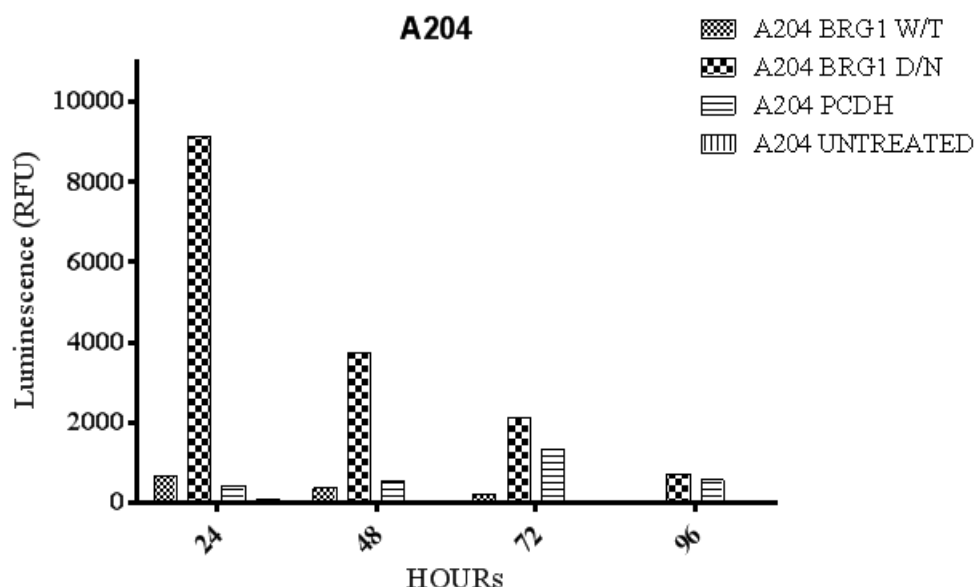


*Figure 4.15 Effect of SMARCA4 ATPase activity inhibition on survival (growth) in G401 and A204 Rhabdoid cells. SMARCA4 D/N causes drastic growth arrest within 36 hours post infection, compared to the SMARCA4 WT and control-vector infectants. Growth was measured using MTT assay absorbance measurements were taken between (76h-168hours) post infection. Results of three individual experiments are indicated. Untreated cells (Un) and empty vector infected cell (PCDH) have been used as negative controls.*

In order, to understand the mechanism of cell growth arrest caused by SMARCA4 inhibition, an apoptotic assay against caspases III and VII was performed. Caspases III and VII are cleaved directly by caspase VIII or by the intrinsic BID (a BH3 domain-containing proapoptotic Bcl2 family member,) pathway (McIlwain, Berger et al. 2013). In accordance with the change in cell morphology, at 36 hour post infection in both G401 and A204 cell lines we detected an increase in apoptotic response; however this response decreases after 48 hours after *SMARCA4* D/N induction, in agreement with the rapid loss of cell viability (Figure 4.16 and Figure 4.17) In particular caspase cleavage is triggered by both *SMARCA4* WT and DN forms although cleavage is twice as great in the DN form. This returns to levels similar to untreated by 96 hours for the WT form but not for *SMARCA4* DN.



*Figure 4.16 Effect of SMARCA4 expression on apoptosis in the G401 cell line. Expression of SMARCA4 D/N induces strong early apoptosis which compromises cell viability. Caspase cleavage was measured by luminescent measurements were taken (24h-96hours) post infection. Graphs show results of three individual experiments. Untreated cells and empty vector infected cell (PCDH) were used as negative controls.*



*Figure 4.17 Effect of SMARCA4 inhibition on apoptosis in the A204 cell line. Expression of SMARCA4 D/N induces strong early apoptosis which compromises cell viability. Expression of SMARCA4 D/N induces strong early apoptosis which compromises cell viability. Caspase cleavage was measured luminescent measurements were taken (24h-96hours) post infection. Graphs show results of three individual experiments. Untreated cells and empty vector infected cell (PCDH) were used as negative controls.*

After *SMARCA4* D/N induction cells appeared visibly deteriorated after 36 hours post infection: DNA and RNA extraction was successful at only this time point. Nevertheless mRNA levels of *CDKN2A*, *CDKN1A* and *CD44* in A204 and G401 cell lines were measured by Q-PCR and reveal that inhibition of *SMARCA4* ATPase activity results in an up-regulation only of *CDKN1A*, reflecting previous reports (Kang, Cui et al. 2004). However, *SMARCA4* D/N expression does not affect the expression of *CDKN2A* and *CD44*, suggesting that *SMARCA4* regulates transcription in a different a manner to *SMARCB1* (Figure 4.18). Interesting also the WT form induces increase of *CDKN2A* and *CDKN1A*; this could be due to early detection of the mRNA level, while the cells were still under the effect of the Puromycin selective agent.

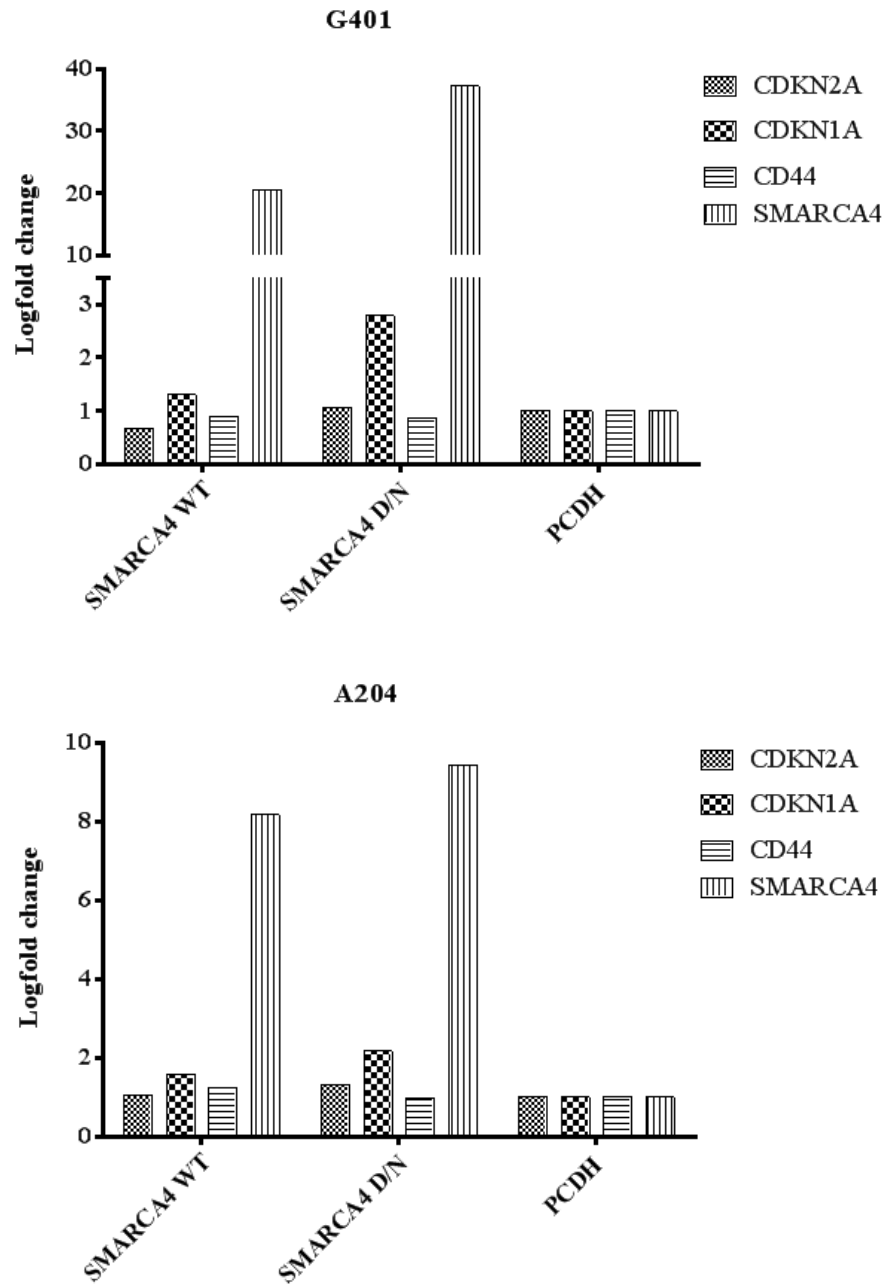


Figure 4.18 Expression of SMARCA4 induces change in expression of CDKN1A, but not CDKN2A and CD44 in G401 and A204 cells. RNA was extracted 36 hours day post infection. The mRNA levels were measured by real-time qPCR and normalised to GAPDH. Relative expression values are the mean of 3 independent measurements normalised to empty vector infected cell (PCDH) which was used as a control.



## 4.6 Discussion

### 4.6.1 Characterisation of the effects of *SMARCB1* loss in Malignant Rhabdoid cells

Earlier studies showed that re-expression of *SMARCB1* in Rhabdoid cell lines causes arrest of cell growth by SWI/SNF recruitment and activation of *CDKN2A* and *CDKN1A* in early stage. For the first time characterisation of the late effects of *SMARCB1* re-expression has been attempted in four cell lines, derived from different location of the body.

A preliminary microarray time course study was performed on a G401 cell line in which *SMARCB1* has been re-expressed. Hundreds of genes were deregulated in response to *SMARCB1* induction; in particular *SMARCB1*'s effect on gene expression gradually increased until the 7<sup>th</sup> day post infection, when a stable steady-state was reached. Based on these findings the 7<sup>th</sup> day post infection was set as a time point for our experiment, aiming to investigate the late effect of *SMARCB1* re-expression.

In this study *SMARCB1* induction was strategically performed in four different cell lines. Cell lines (G401, A204, STA-WT1 and CHLA-266) derived from different locations of the body, reflecting the complexity of the Rhabdoid Tumours. *SMARCB1* re-expression in Rhabdoid cells causes a reversion of the cells tumourigenic status into one of permanent senescence: cells showed characteristic changes in morphology such as enlarged and flattened cytoplasmic shape and increased granularity. This is concomitant, in part, to a persistent up-regulation *CDKN2A*, *CDKN1A* and *CD44* and therefore to an arrest of cell growth. These effects were detectable in all four Rhabdoid cell lines (with the exception that *CD44* is not expressed by CHLA-266), regardless of the original tumour site of origin.

*SMARCB1* encodes the SMARCB1 subunit of a chromatin remodelling complex (SWI/SNF). Here we demonstrated SMARCB1 function plays an essential role in

recruitment of the SWI/SNF complex at chromatin level. In particular formaldehyde cross-linking and chromatin immunoprecipitation (ChIP) were used to investigate occupancy of the *CDKN2A* and *CDKN1A* promoters in living Rhabdoid tumour cells. *SMARCB1* re-expression persistently triggers activation of *CDKN2A* and *CDKN1A* promoters concomitant SMARCA4 recruitment to their respective promoters, accompanied by increase of H3K4me (gene activation marker) and decrease of H3K27me (repressive marker).

In conclusion, this study demonstrated that *SMARCB1* re-expression alone induced senescence in Rhabdoid cell lines, by a lasting up-regulation of *CDKN2A* and *CDKN1A* and CD44 through direct recruitment of SMARCA4 and H4k4me3 at the TSS of these genes.

#### **4.6.2 Effects of *SMARCB1* loss on methylation in Malignant Rhabdoid cells**

Demethylating agent 5-azaCdR has been extendedly used to elucidate the implication of methylation in gene expression and tumourigenesis. Recently, a study in glioblastoma patients suggests a correlation between *SMARCB1* loss and change in methylation status (Zunarelli, Bigiani et al. 2011) moreover, ChIP analyses show a direct effect of *SMARCB1* in methylation modifications at the histone level and in recruiting SWI/SNF complexes at the promoter region of *CDKN2A* and *CDKN1A*.

In this study Rhabdoid cell lines were treated with 5-azaCdR to investigate *SMARCB1* involvements in the DNA methylation process. Transcriptional activation of *DAZL* (embryonic germ cell marker), *RASSF1A* (Ras associate oncogene) and *HTATIP2* (metastasis suppressor gene) was observed followed 5-azaCdR treatment in a variety of paediatric cancer cell lines, including Medulloblastoma (Harada, Toyooka et al. 2002). These findings suggest that aberrant methylation at the promoter of *DAZL*, *RASSF1A* and *HTATIP2* promotes tumourigenesis in Rhabdoid Tumours

DNA methylation and histone lysine methylation are one of a number of processes contributing to transcriptional regulation. Methylation of H3k4 appears to inhibit methylation at CpG Islands (Weber, Hellmann et al. 2007), as confirmed by the methylation at histone H3k4 in primary human leukaemia cells, following 5-azaCdR treatment (Zou, Ma et al. 2012). In Rhabdoid cell lines 5-azaCdR treatment induced up-regulation of *CDKN2A*, *CDKN1A* and *CD44*, in a similarly manner but with a lesser magnitude to that observed following *SMARCB1* re-expression. These observations validate the hypothesis that hSNF5 and SWI/SNF complexes regulate gene expression by direct modulation of DNA- methylation.

#### **4.6.3 Tumourigenesis induced by *SMARCB1* loss is affected by aberrant activity of SMARCA4 ATPse subunit**

The SMARCA4 ATPse subunit has been found mutated or silenced in a variety of tumours and mutations of the *SMARCA4* gene are more common than *SMARCB1* (Shain and Pollack 2013). The presence of germline nonsense mutations and/or somatic inactivation of *SMARCA4* in Rhabdoid patients have also provided evidence of SMARCA4 involvement in Rhabdoid tumourigenesis (Schneppenheim, Fruhwald et al. 2010, Hasselblatt, Gesk et al. 2011). *SMARCA4*'s role in Rhabdoid Tumours has previously been partially demonstrated in G401, A204 and *SMARCB1*-conditional mice, in which siRNA knock down of SMARCA4 induced growth arrest in G401 cell and blocked tumour formation in mice (Wang, Sansam et al. 2009).

A dominant/negative and a wild type SMARCA4 were stably induced in G401 and A204 Rhabdoid cells. Disruption of SMARCA4 activity causes a drastic arrest of cell growth, inducing apoptosis. In response to cleavage of caspase III and VII the cells undergo apoptosis just 48 hours post infection, showing cell shrinkage, nuclear fragmentation and chromatin condensation. SMARCA4 D/N expression in Rhabdoid cells also leads to an up-regulation of *CDKN2A*, while *CDKN1A* and *CD44* remain unchanged.

In conclusion, stable inactivation of SMARCA4 ATPase activity causes cell growth arrest in part by induction of apoptosis. These findings validate the hypothesis of an aberrant activity of SMARCA4, following *SMARCB1* deletion. However, this study also demonstrates that these two SWI/SNF complexes, even though they are part of the same protein complex and the same tumourigenic process acts upon different genes and therefore different pathways.

## **Chapter 5      Genomic and Methylation Analysis of Malignant Rhabdoid Tumours**

## 5.1 Introduction

In the last decade DNA and RNA microarrays have provided invaluable information on biological mechanisms, resulting in understanding of pathogenesis, in characterisation of biomarkers for diagnosis and in the prediction of therapeutic responses. However, these technologies present various limitations, overcome by the development of next-generation sequencing (NGS) technologies and more sophisticated and precise array methods. NGS provides an unprecedented resolution and throughput which allows the exploitation of entire genome and transcriptome sequences. Whole-genome sequencing has been extensively carried out to identify genomic alterations, gene fusions, somatic mutations and alternatively spliced forms in colon, lung, breast, prostate and pancreatic cancers, Ovarian Clear Cell Carcinoma, Hepatocellular Carcinoma, Leukemia and Melanoma (Ku, Cooper et al. 2013). Aberrant DNA methylation is a well recognised hallmark of most cancers. New generations of methylation arrays have been used to interrogate methylation sites across the genome at single-nucleotide resolution. A comprehensive view of the methylation landscape has been evaluated in Choroid Plexus Tumours, Lymphoblastic Leukaemia, Medulloblastoma, Paediatric Glioblastoma, characterising distinct methylomic events for each cancer group (Castelo-Branco, Choufani et al. 2013).

Although previous studies have comprehensively evaluated molecular profiling of primary cancer such as Medulloblastoma, few studies were performed in Malignant Rhabdoid Tumours (Ma, Kao et al. 2010, Sullivan, Yankovich et al. 2012, Takita, Chen et al. 2014). Moreover, RNA expression profiling of these malignancies were performed just by Affymetrix arrays, while methylation profiling has never been published. In the past few years just two studies have attempted to investigate the overall differential gene expression of Rhabdoid Tumours: these studies characterised a limited number of significantly deregulated genes in Rhabdoid Tumours when compared with other paediatric malignancy and identified gene signatures (down-regulation of genes associated with PRC occupancy and genes associated with chromatin processing/remodeling) that might distinguish Extra Cranial Rhabdoid (ECRT) from Atypical Teratoid Rhabdoid Tumours (ATRT) (Birks, Donson et al. 2013). However

these analysis are limited by the number of primary samples and the technique used: samples were were processed by different laboratories and analysed using different platforms.

In this study, for the first time we collected 39 primary Rhabdoid samples (13 Extra Cranial Rhabdoid Tumours and 26 Atypical Teratoid Rhabdoid Tumours) and analysed them using Genome-wide approaches. In particular, this is the first study to characterise global transcriptional and methylomic changes in paediatric cancer using high-throughput techniques. This approach gives a unique opportunity to identify the complex mechanism of tumourigenesis in Rhabdoid Tumours and to provide a catalogue of tumourigenic events that can be further investigated.

## 5.2 Aims

The genome-wide relationship between gene methylation and expression in Malignant Rhabdoid Tumours has not previously been investigated. Using global DNA methylation and whole transcriptome studies on primary Rhabdoid Tumours, this study aims:

- To characterise differential patterns of DNA methylation and gene expression, and their roles in development of Malignant Rhabdoid Tumours
- To investigate possible methylation-dependent gene expression alterations dependent on *SMARCB1*
- To identify possible subgroups with specific methylation or gene expression events correlating with clinico-pathological features of primary tumours.



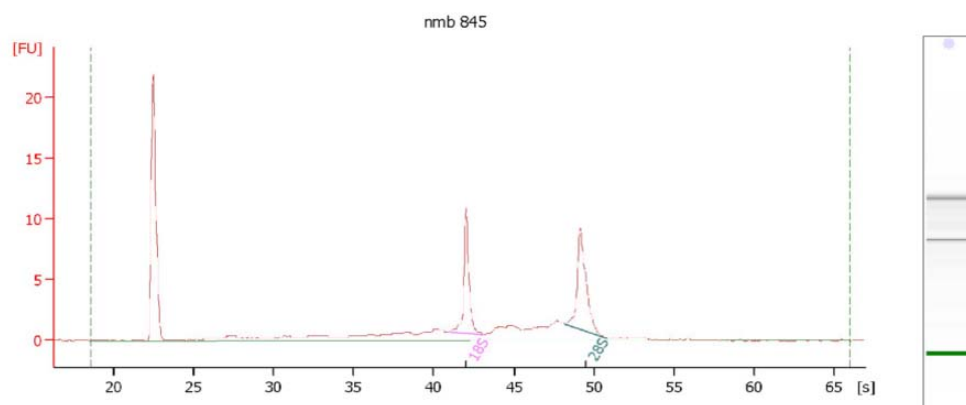
### **5.3 RNAseq analysis identifies Malignant Rhabdoid Tumours as a single biological entity containing subgroups**

In this section a genomic analysis was designed to test the hypothesis that patterns of gene expression were shared within the Rhabdoid Tumours , regardless of the location of the malignancies. In this analysis 23 Rhabdoid primary tumours (11 Atypical Teratoid Rhabdoid Tumours and 12 Extra Cranial Rhabdoid Tumours) were compared with 251 Medulloblastoma samples to characterise transcriptomic and methylation diversity within the two paediatric malignancies. Medulloblastoma was chosen as an outgroup control primarily because a large volume of comparable RNA-seq data was available and because it is another paediatric embryonal tumour.

#### **5.3.1 Primary samples preparation**

23 *SMARCB1* negative Malignant Rhabdoid Tumour (MRT) primary samples were collected from multiple centres throughout the UK in collaboration with the CCLG bio bank ([www.cclg.org.uk/tissue-bank](http://www.cclg.org.uk/tissue-bank)). All samples were evaluated at the referring centre and subsequently centrally review by the CCLG bio bank. RNA was extracted from frozen sample following the previously described procedure (section1.1.1).

RNA integrity number (RIN score) was evaluated for each sample, using a Bioanalyzer. All RNA samples were scored with a RIN greater than 7. Figure 5.1 shows an example of the Bioanalyzer trace (sample nmb 845)



**Overall Results for sample 1 : nmb 845**

RNA Area:	57.4	RNA Integrity Number (RIN):	8.3 (B.02.08)
RNA Concentration:	34 ng/μl	Result Flagging Color:	<div style="background-color: #ccccff; border: 1px solid black; width: 50px; height: 15px;"></div>
rRNA Ratio [28s / 18s]:	1.2	Result Flagging Label:	RIN: 8.30

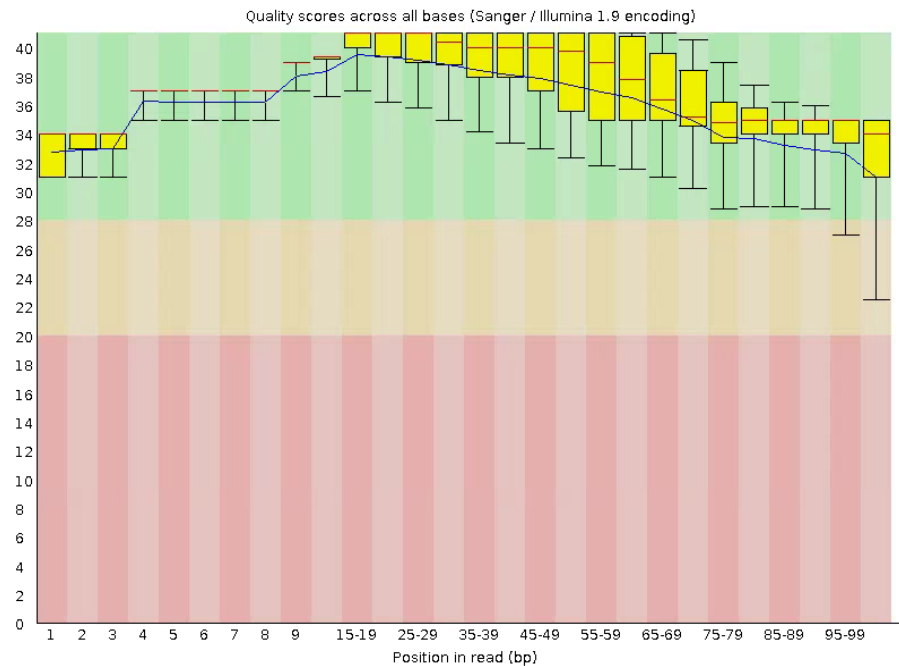
**Fragment table for sample 1 : nmb 845**

Name	Start Time [s]	End Time [s]	Area	% of total Area
18S	40.94	43.09	9.1	15.9
28S	48.15	50.78	10.8	18.9

*Figure 5.1 Electropherogram of sample nmb845 with 8.3 RIN score*

### 5.3.2 RNA-seq QC analysis

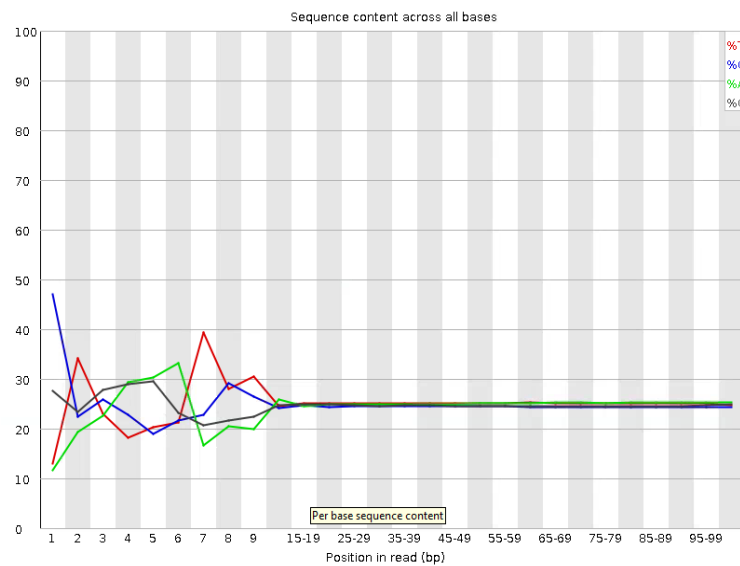
All samples submitted for sequencing were considered to have passed quality control checks as measured by FASTQC. For instance, 100% of samples presented a Phred score across all bases higher than 30 (Figure 5.2).



*Figure 5.2 Per base sequence quality derived from sample nmb 845 from a FASTQC html report. The red line represents the median value, while the blue line the quality. The plot presents 3 different background shading: green indicates good quality score.*

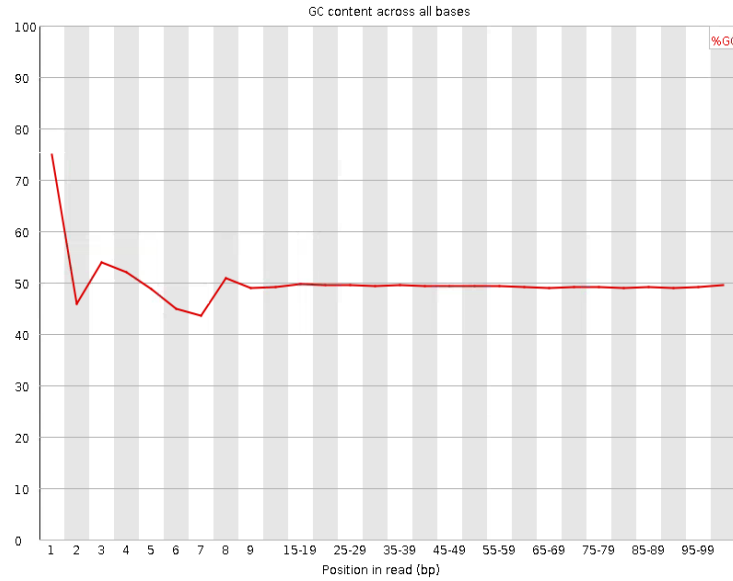
3 of the 10 examined quality modules in the FASTQC analysis presented warnings: Per base sequence content, Per base GC content and Sequence Duplication levels however this is typical of the platform and library preparation used every sample analysed regardless of type shows these issues and should not be considered problematic. In Per base sequence content the proportion of the four DNA bases for each base position in a

sequence is calculated. For per base GC the proportion is considered only for the G and C content (Figure 5.4 and Figure 5.4 ). This ratio is expected to be almost zero between the different bases of a sequence run, however in a practical sense high variability in the first 10 bases is often observed as result of non-random priming (Hansen, Brenner et al. 2010) but does not in itself imply false sequence.

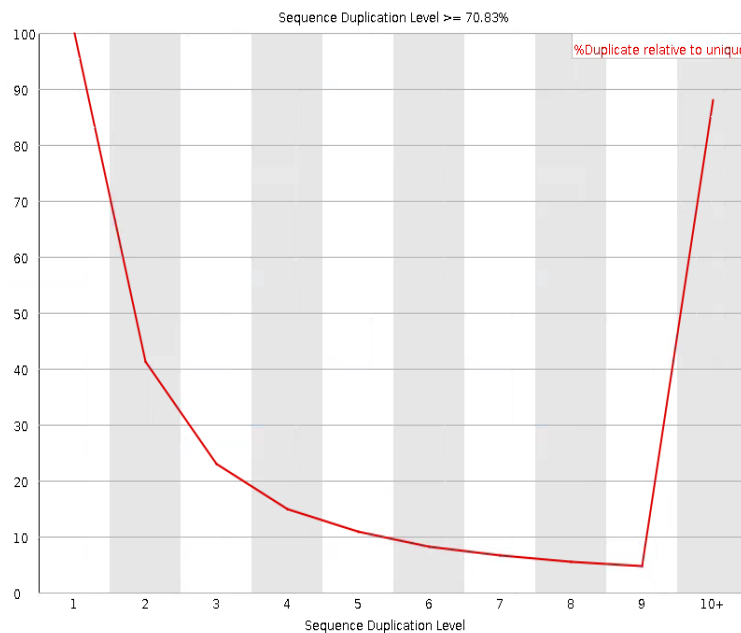


*Figure 5.3 Per base sequence content analysis of sample nmb843 from FASTQC html report. High variability of the first 10 bases was observed in all samples of our data set; this phenomenon is more plausibly constitutive biases rather than batch effect.*

The sequence duplication levels module evaluates the first 50 cycles to estimate the degree of duplication for every sequence in the data set. In our cases approximately 25% of the analysed sequences presented a degree of duplication higher than 10, consistent with the hypothesis of biases due to random priming (Figure 5.5).



*Figure 5.4 Per base GC content of sample mnb 843 from FASTQC html report. High variability of the first 10 bases was observed in all samples of our data set; this phenomenon is more plausible to be constitutive biases rather than batch effect.*



*Figure 5.5 Sequence duplication levels plot of sample mnb483 from FASTQC html report. The 10th column includes all duplication levels greater or equal 10.*

### 5.3.3 QC analysis

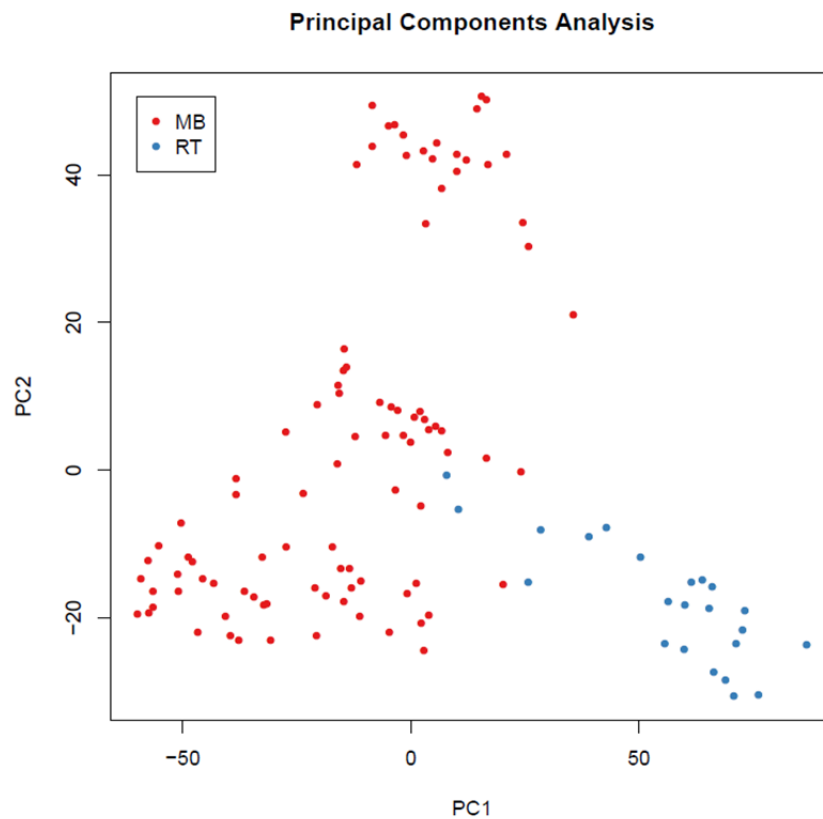
#### 5.3.3.1 *Unsupervised analysis*

\*.Bam files were converted to read counts using the python script HTseq-count (as detailed in section 4.2.2.2) to facilitate accurate comparisons between samples. Normalisation consists of dividing the gene counts by the total number of reads in each library. The resulting data was then subject to a log2 variance stabilising transform (vst) using the DE-seq package (R/Bioconductor) and then subjected to unsupervised clustering approaches (PCA, HC, and NMF), attempting to detect true biological differences between Rhabdoid and Medulloblastoma primary tumours. The analysis was conducted considering all the Medulloblastoma samples for the NMF, GSEA and DAVID, however for clarity of visualisation PCA analysis and heatmap were computed on a smaller subset of 92 primary tumours.

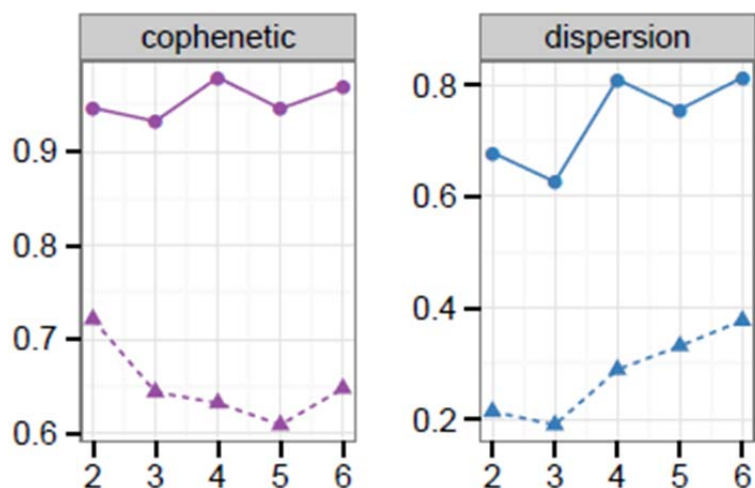
Principal component analysis (PCA) was carried out with a subset of 5377 gene transcripts these were filtered as those genes showing high variance and expression; i.e with all samples showing a vst score greater than 5, an interquartile range greater than 4 and/or an overall range greater than 6 (prop = 0.2, A=5, Gap = 6, interquartile range IQR = 4). The PCA plot shows the tendency of sample data to form two different groups tending to cluster the Rhabdoid group separate from Medulloblastoma (Figure 5.6).

NMF/consensus clustering was used to identify the optimal number of metagenes and clusters (see section 3.34). The stability of the clustering was evaluated by considering cophenetic correlation coefficient and dispersion coefficient. Cophenetic correlation is defined as the Pearson correlation between the distances of sample induced by the consensus matrix and their cophenetic distances from a hierarchical clustering based on these distances (Brunet, Tamayo et al. 2004); dispersion coefficient defines the general quality of the clustering consistency (Kim and Park 2007). Metagenes were computed using the above filtered genes and showed a cophenetic correlation coefficient value equal to 1 with a 4 and a 5 metagene solution, while dispersion coefficient indicated an

increase directly proportional to the number of the metagenes ( Figure 5.7). Two robust consensus clusters were identified for  $k=4$ .



*Figure 5.6 Unsupervised Principal Component Analysis of DEseq data set distinguishes two different biological groups. In red are represent Medulloblastoma primary sample while in blue the Rhabdoid. The x-axis is the first component; the y-axis is the second principal component.*



*Figure 5.7 Visualisation of cophenetic correlation and dispersion measures to assess the stability of clustering associated with each rank  $k$  for the Rhabdoid and Medulloblastoma derived DEseq data. Dotted line indicates randomised data and non-dotted actual data.*

We performed the unsupervised NMF analysis ( $k=4$ ) in the gene subset for all samples to determine if there was evident structure in the expression information that reflected the tumour phenotype. Analysis of the data revealed clear patterns of gene expression that separated the Rhabdoid and Medulloblastoma Tumours (subdivided into 3 metagenes: corresponding to WNT, SHH and a third subgroup). The HC algorithm also confirmed the clustered subgroups of samples. To evaluate the reproducibility of this analysis, the data were randomly re-clustered 100 times (McShane, Radmacher et al. 2002). The consensus analysis reconfirmed the 5 subgroups (Figure 5.9).

The heatmap and box plot indicated in Figure 5.9 contains the top 20 differentially expressed genes (defined by p-value) and illustrates the consistency of the separation between Medulloblastoma and Rhabdoid Tumours.



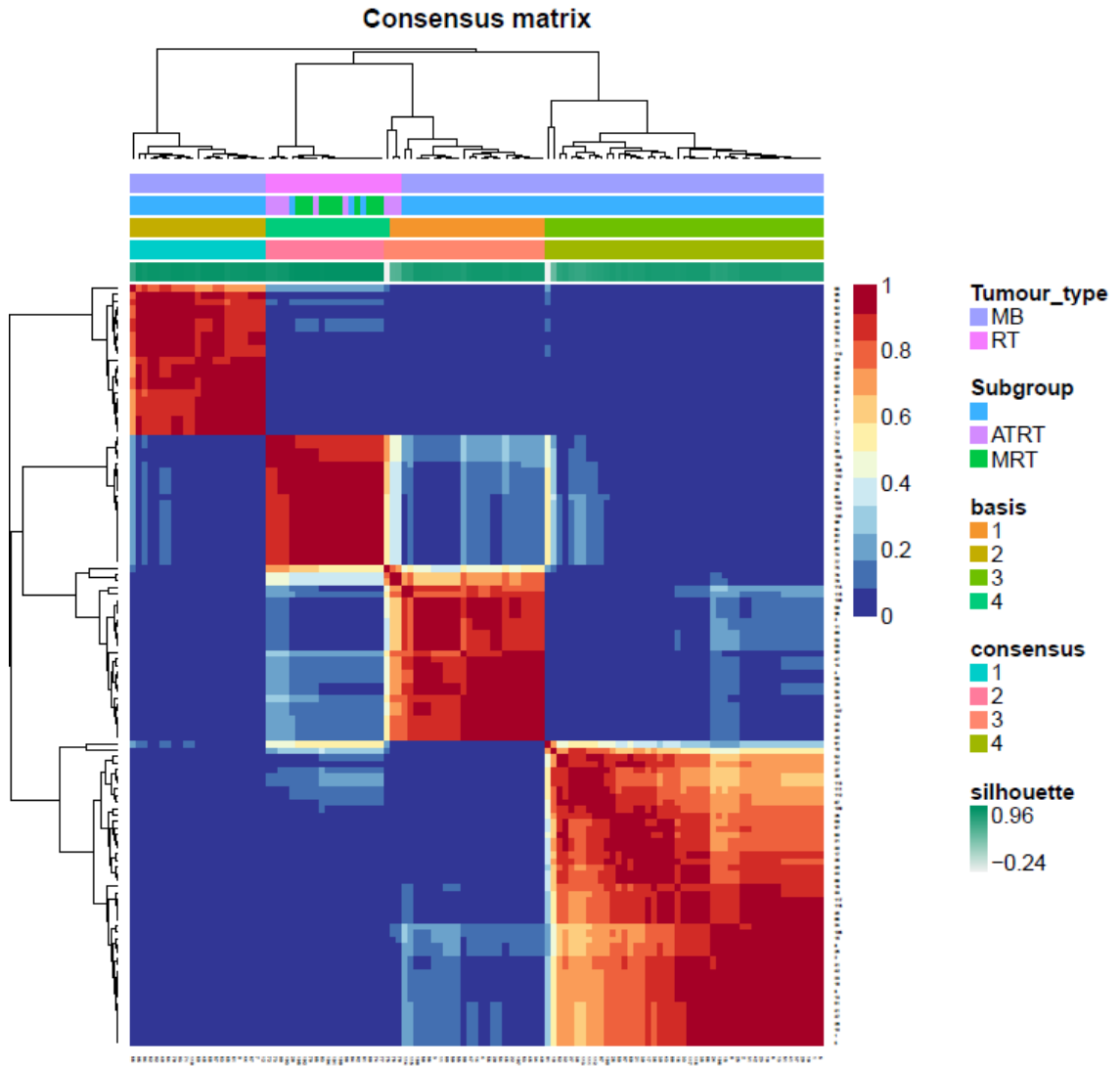


Figure 5.8 Unsupervised NMF and hierarchical cluster analysis ( $K=4$ ) on DEseq data set characterises four different metagene classes. Rhabdoid and Medulloblastoma are two different groups, the genetic complexity of Medulloblastoma (characterized by 4 different subgroups: WNT, SHH, and Group 3 and Group 4 clustered together) are retained in this analysis. The matrices were coloured 0 (deep blue, indicating samples never in the same cluster) to 1 (dark red, indicating samples always in the same cluster).

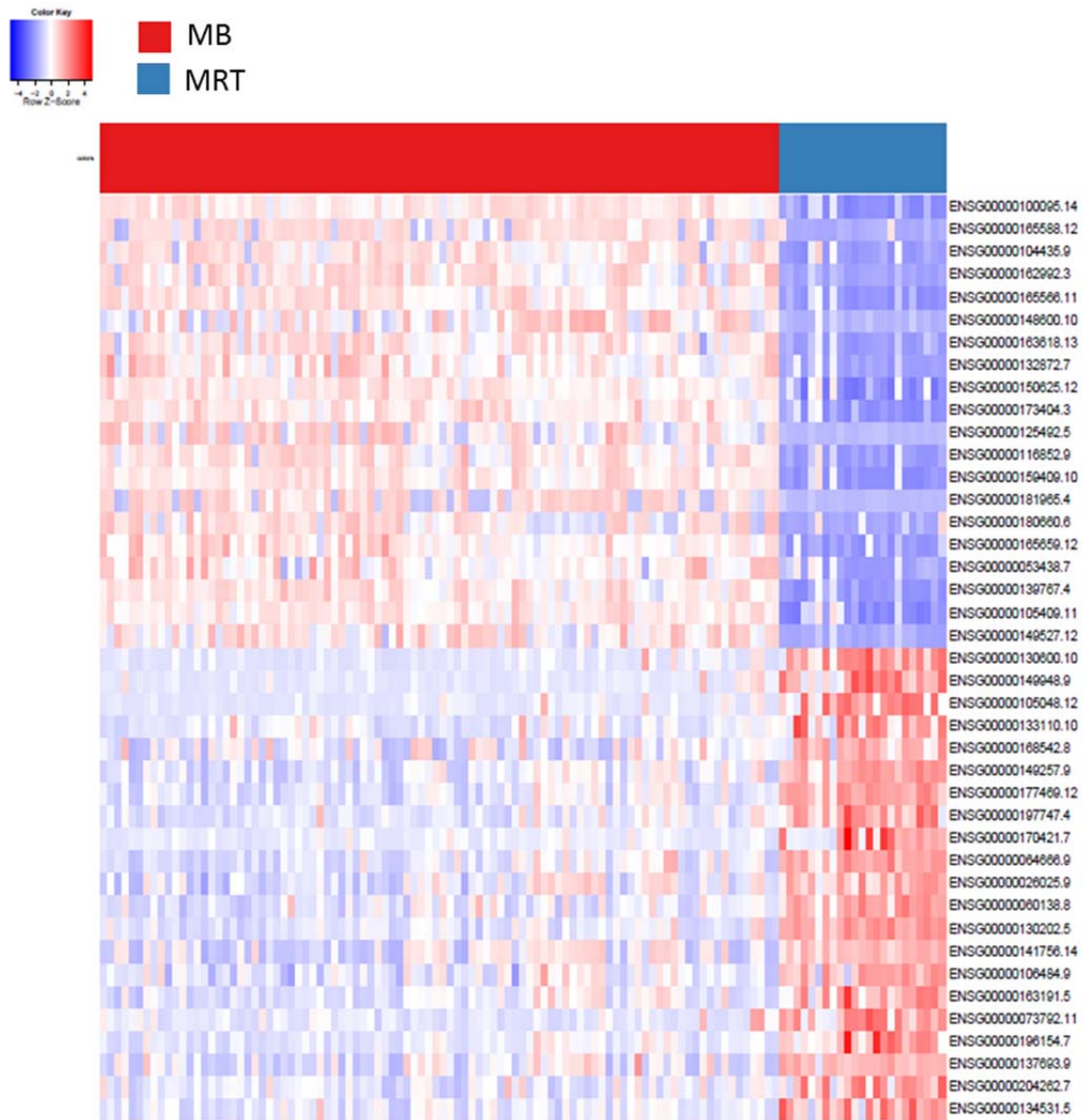


Figure 5.9 Heatmap of gene expression showing the 20 most significantly differentially regulated genes between Medulloblastoma and Rhabdoid Tumours as calculated by adjusted p-value from DESeq2. Each row represents a gene and each column a sample. Overexpressed (pale red to dark red) and underexpressed transcripts (pale blue to dark blue).

#### 5.3.3.2 *Supervised analysis*

Raw count values were analysed in a pair wise fashion using DESeq2 in R. DESeq2 identifies differentially expressed genes using a statistical method based on the negative binomial distribution (Anders & Huber 2010). It computes False Discovery Rate (FDR), folds change and adjusted p-values which are then subjected to the Benjamini-Hochberg correction for multiple testing. The data set was first tested for variance dispersion, p-value distribution and FDR. Estimation of the data variance is one of the most important steps in the differential expression analysis. In particular, this test allows to distinguish noise due to technical error and evaluate if variation observed are real biological effects and not a result of experimental conditions. Figure 5.1 represents the estimate of dispersion in our data set, where the black dots represent the empirical dispersion values and the red line the fitted values. The means across the genes were equally distributed, as indicated by the dispersion trend. DESeq computes also a preliminary differential gene expression significance test, calculating the significance of change in expression of gene between two conditions (in this case Rhabdoid and Medulloblastoma primary tumours). This test returns adjusted  $p$ -values, mean count values and log2 fold change. The analysis is visualised using an MA plot where log2 fold values for each gene are plot against the means and the significant differential expression genes at 10% FDR are coloured in red (Figure 5.10, Figure 5.11).

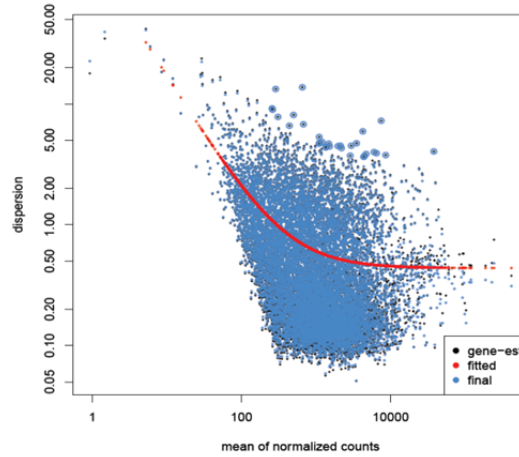


Figure 5.10 Gene expression dispersion estimation and model fitting of Rhabdoid and Medulloblastoma derived DEseq data. Each dot represents a per gene dispersion estimation and the red line the fitted mean dispersion function.

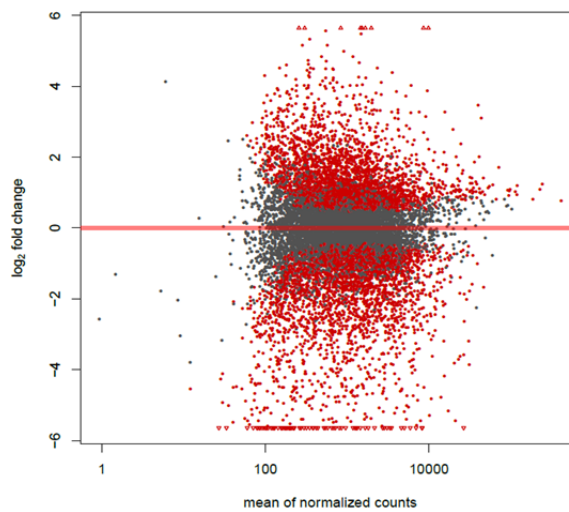
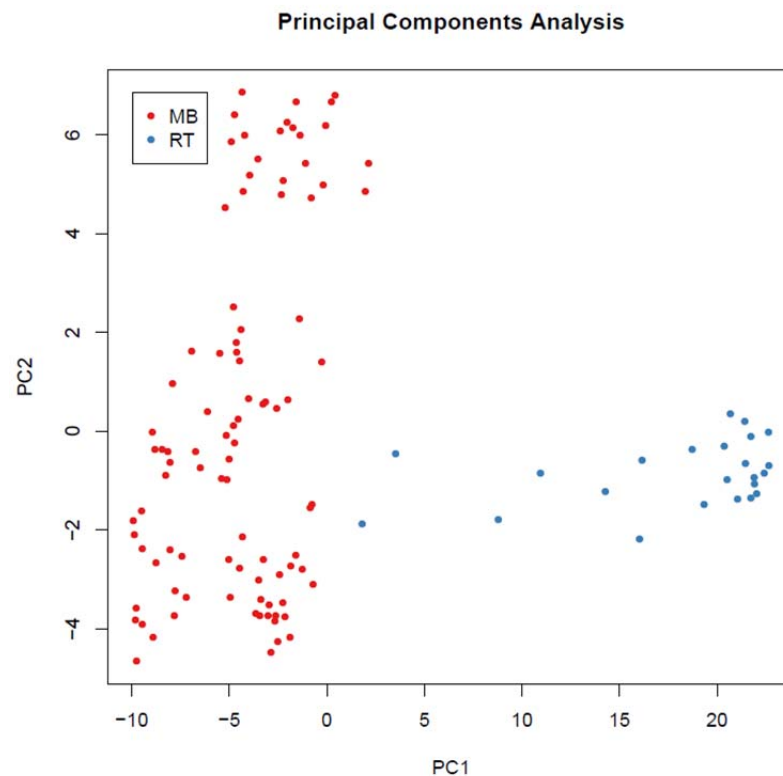
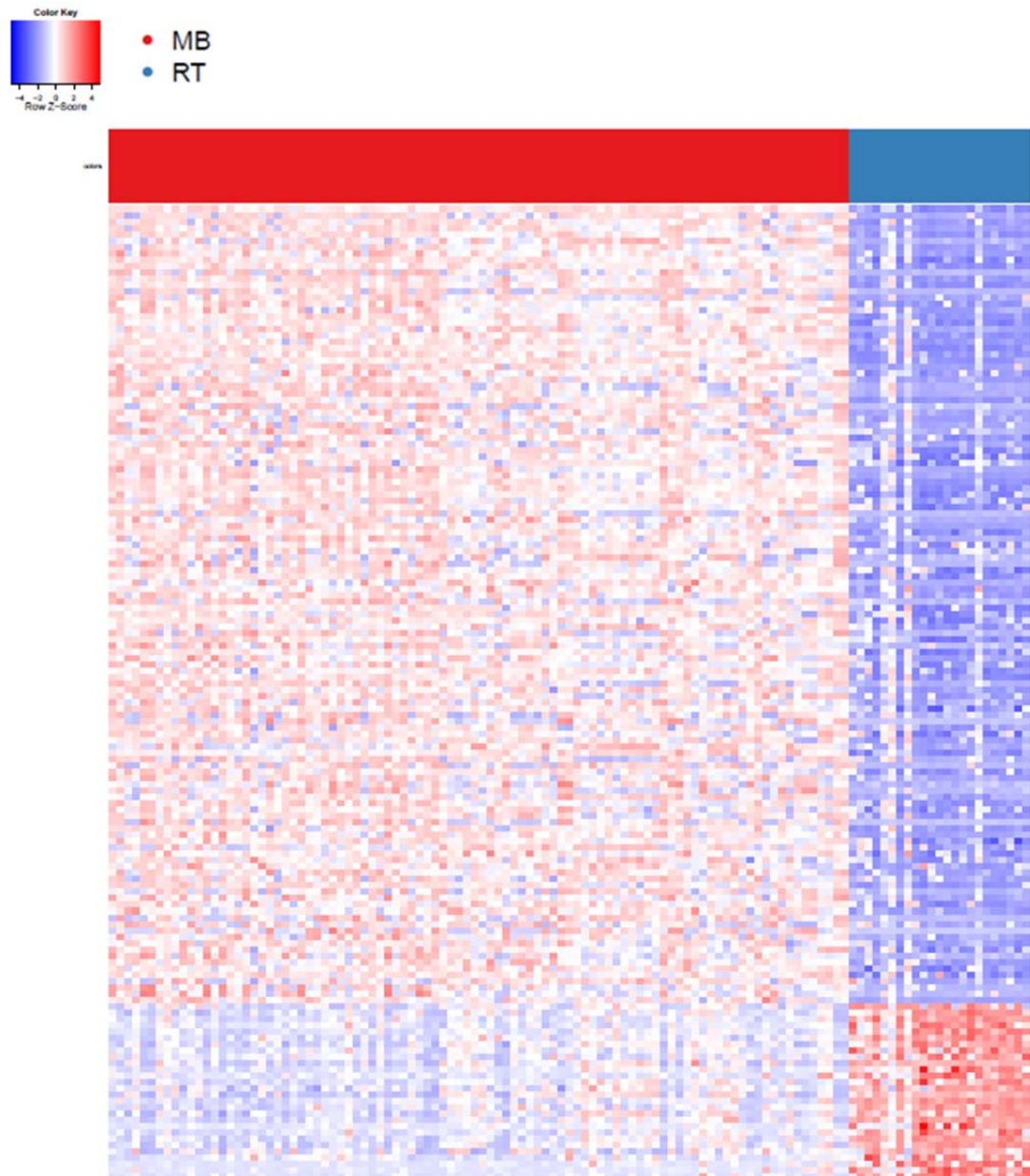


Figure 5.11 Per gene dispersion estimates from Rhabdoid and Medulloblastoma derived DEseq data. Differentially expressed genes Log<sub>2</sub>fold versus mean expression over all samples. Red dots represent genes marked as differentially expressed at 10% FDR, when Benjamini and Hochberg method is applied. The lines on the top and on the bottom of the plot represent genes with a very low or very high log fold change (i.e. the value off scale).

Next a supervised cluster analysis was performed. Supervised clustering classifies the objects in order to capture between-class variations with *prior* knowledge on data clusters previously obtain (Dettling and Buhlmann 2002). First the DEseq data were filtered in order to obtain a data set composed of 157 significant genes (adjusted *p*-value < 0.001; moderated log fold change  $\leq -3$  and  $\geq +3$ ). PCA analysis was performed confirming the ability of these 157 genes to distinguish between Rhabdoid and Medulloblastoma primary tumours (Figure 5.12). Filtered data were used to generate a supervised heatmap which visualises distinct gene expression of Rhabdoid and Medulloblastoma primary groups (Figure 5.13).



*Figure 5.12 Supervised principal component analysis of 157 genes selected as significantly differentially expressed between Rhabdoid Tumours and Medulloblastoma distinguishes two different biological groups. In red are represented Medulloblastoma primary sample while in blue the Rhabdoid. The x-axis is the first component; the y-axis is the second principal component.*



*Figure 5.13 Supervised heatmap of gene expression in Medulloblastoma and Rhabdoid Tumours. Overrepresented (pale red to dark red) and underrepresented transcripts (pale blue to dark blue). Each row represents a gene while each column represents a sample. Overexpressed (pale red to dark red) and underexpressed transcripts (pale blue to dark blue).*

The DEseq results were filtered for values with an adjusted  $p$ -value greater than 0.0001 and a moderated log2 fold change less than -4 or greater than 4. In Table 5.1 are indicated the 20 most differentially expressed genes as a result of filtering, relative difference in expression of these genes is indicated in Figure 5.14 and Figure 5.15. Twelve were associated with tumour progression (\*), two involved in cell proliferation (+), one directly involved in tumour invasion and angiogenesis (-), one indicated as potent tumour suppressor genes (~) and one acting in neurodevelopment proliferation of neural progenitor cells (#) ([www.genecards.org](http://www.genecards.org)). Complete list in Appendix I.

Gene Symbol	Adjusted <i>p</i> -value	Moderate log2fold
<b>KRT8*</b>	1.31E-51	2.691206
<b>AMER2*</b>	7.10E-37	2.446738
<b>HMGA2*</b>	4.96E-26	2.001177
<b>TNNT1 +</b>	1.48E-23	1.617837
<b>POSTN *</b>	6.01E-34	1.59003
<b>INSM1 *</b>	2.14E-33	1.582486
<b>S100A4 -</b>	9.74E-48	-1.700879104
<b>KIF21B</b>	1.50E-37	-1.728031529
<b>ST18 ~</b>	2.52E-29	-1.744469767
<b>SRRM4</b>	1.06E-34	-1.756456049
<b>COL5A2*</b>	4.18E-39	-1.78873522
<b>NEUROD1 *</b>	1.18E-82	-1.813588525
<b>IGF2BP2 +</b>	3.20E-25	-1.819608749
<b>S100A11 * #</b>	7.95E-35	-1.944165165
<b>IMPG2</b>	1.79E-53	-1.945484655
<b>S100A10*#</b>	1.19E-40	-1.951875084
<b>EPHA8</b>	1.38E-60	-1.958972254
<b>UNCX#</b>	9.58E-71	-2.012249694
<b>NEUROG1*</b>	1.26E-75	-2.024899986
<b>BARHL1 *</b>	7.88E-94	-2.048115239

*Table 5.1 Differentially expressed genes in Rhabdoid vs Medulloblastoma primary tumours. Functional annotation of genes relevant to cancer is shown in the table with \* for genes associated with tumor progression (\*),+ for genes involved in cell proliferation,- for genes directly correlated with tumour invasion and angiogenesis, ~ for genes indicated as potent tumour suppressor genes # for genes acting in neurodevelopment proliferation of neural progenitor cells (#).*



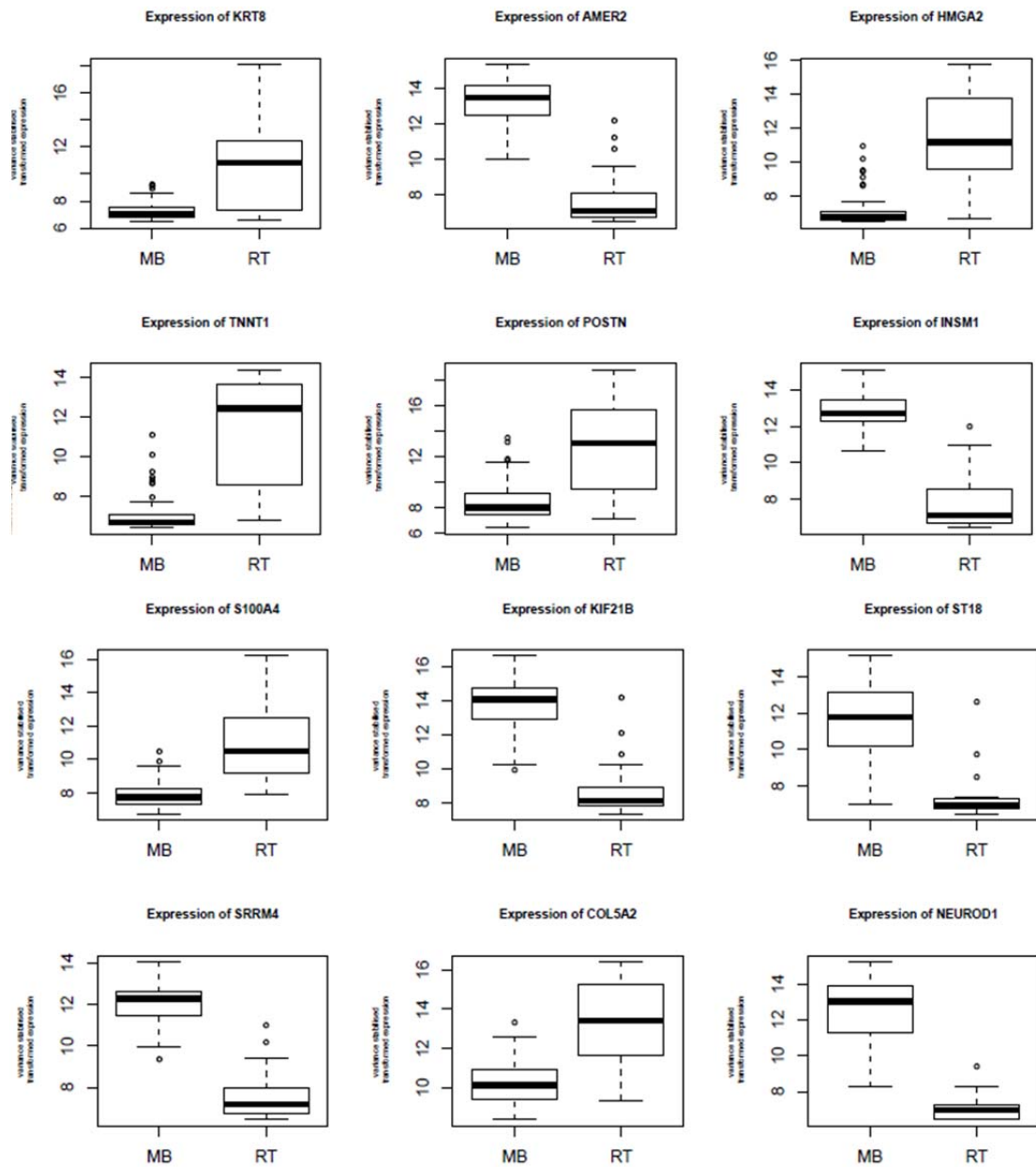


Figure 5.14 Relative difference in expression of 1-12 of most deregulated genes between Rhabdoid Tumours (RT) and Medulloblastomas (MB). Boxplots show Variant stabilised transformed across the samples. Within Boxplots median score (thick black line) and inter-quartile ranges (extent of box) and range (whiskers) are indicated.

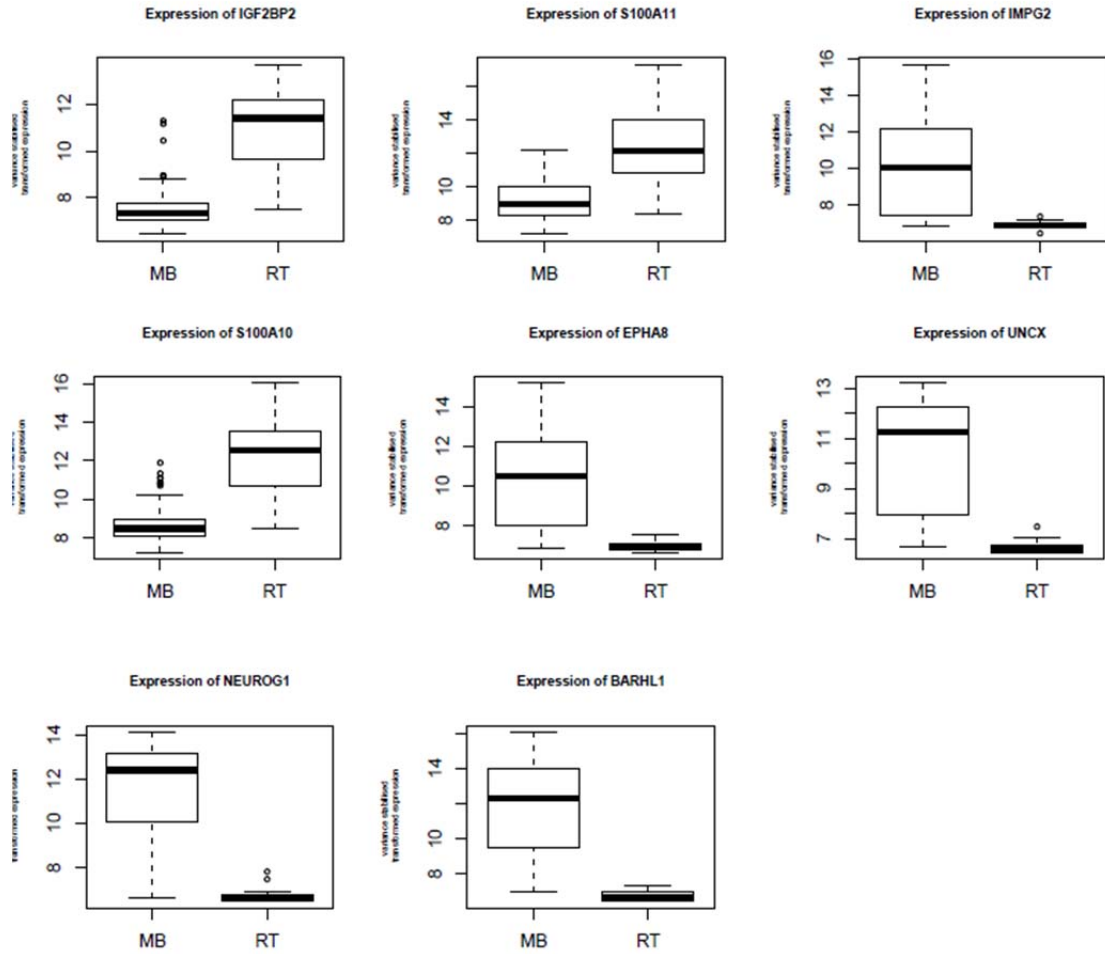


Figure 5.15 Relative difference in expression of 13-20 most deregulated genes between Rhabdoid Tumours (RT) and Medulloblastomas (MB). Boxplots show Variant stabilised transformed across the samples. Within Boxplots median score (thick black line) and inter-quartile ranges (extent of box) and range (whiskers) are indicated.

In particular, *AMER2* (upregulated in Rhabdoid Tumours) has been characterised as a negative regulator of Wnt signalling (Pfister, Tanneberger et al. 2012), while *HMGA2* causes tumourigenesis by enhancing E2F1, by binding to DNA and introducing structural alterations in chromatin. (Fedele, Visone et al. 2006). *STI8* is also indicated as a tumour suppressor gene in breast cancer (Jandrig, Seitz et al. 2004). *NEUROG1* has been found methylated in uveal melanoma together with RASSF1A (Merhavi, Cohen et al. 2007).

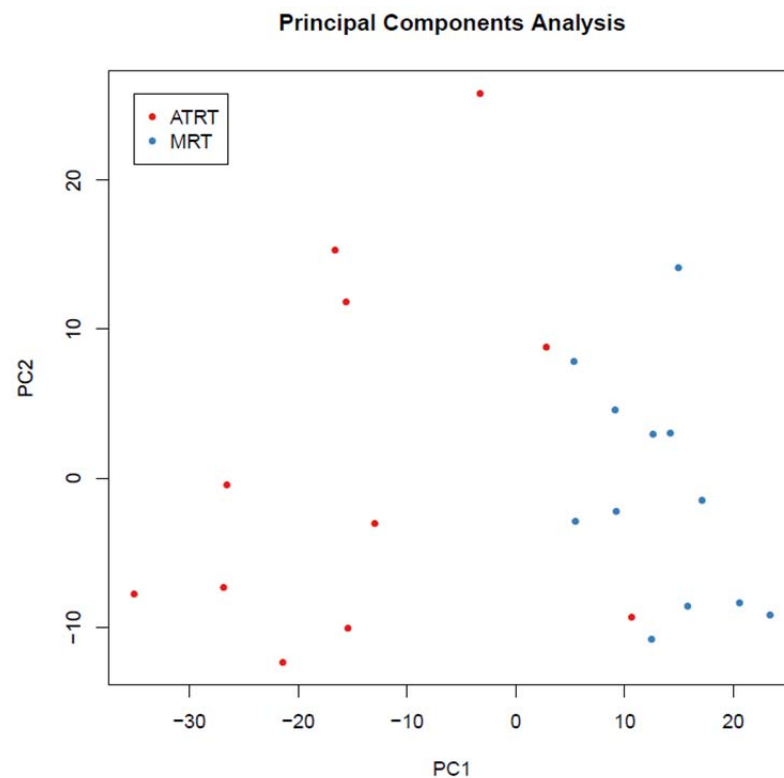
### **5.3.4 Cluster analysis identifies two major subgroup in the Rhabdoid cohort mainly related to location of the malignancies**

In this section we investigated further our Rhabdoid cohort to understand what if any was the significant effect of tumour site on biology. The analysis was first performed unsupervised, as described in the previous section (3.35).

#### *5.3.4.1 Unsupervised analysis*

Principal component analysis (PCA) was carried out with a subset of 1243 genes transcripts (prop = 0.1, A=2, Gap = 3, IQR = 2) from HTseq-count normalised counts transformed into vst data. The PCA plot shows the tendency of sample data to cluster into two separate groups, mainly correlated with the location of the malignancies: Extra Cranial tumours and brain tumours although two Atypical Teratoid Rhabdoid Tumours appear to group with extracranial rhaboid tumours(Figure 5.16).

NMF/consensus clustering approach was used to identify the optimal number of metagenes and clusters (see section 3.34). Metagenes were produced for ranks k=2 to k=5 evaluation of cophenetic and dispersion coefficients indicated that the most robust consensus clustering was k= 2 and 3 (Figure 5.17).



*Figure 5.16 Unsupervised principal component analysis on DEseq data set identifies two major subgroups within the Rhabdoid cohort. In red are represented Atypical Teratoid Rhabdoid Tumours (ATRT) primary sample while in blue the Extra Cranial Rhabdoid Tumours (ECRT). The x-axis is the first component; the y-axis is the second principal component.*

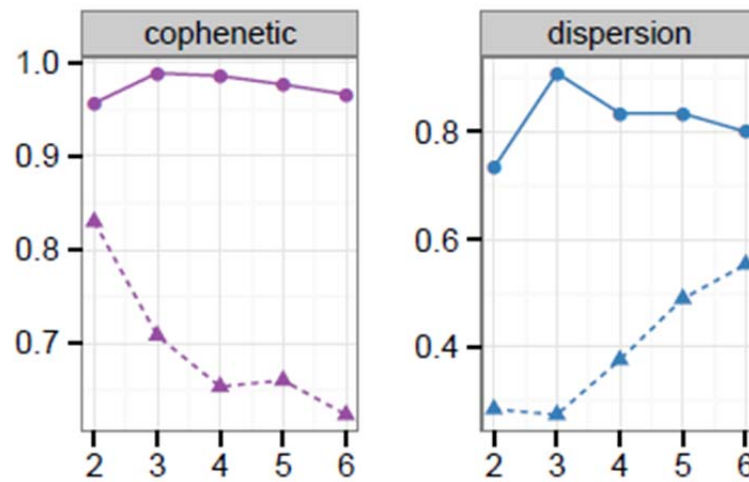


Figure 5.17 Visualisation of cophenetic correlation and dispersion measures to assess the stability of clustering associated with each rank  $k$  for the Rhabdoid derived DEseq data. The dotted line indicates randomised data and non-dotted actual data.

Consensus NMF unsupervised analysis ( $k=2$ ) in the 1243 gene subset reveals two separated patterns of gene expression associated with primary tumour location; in particular one biological group is composed of Extra Cranial Rhabdoid Tumours and the second by Atypical Teratoid Rhabdoid Tumours (Figure 5.18).

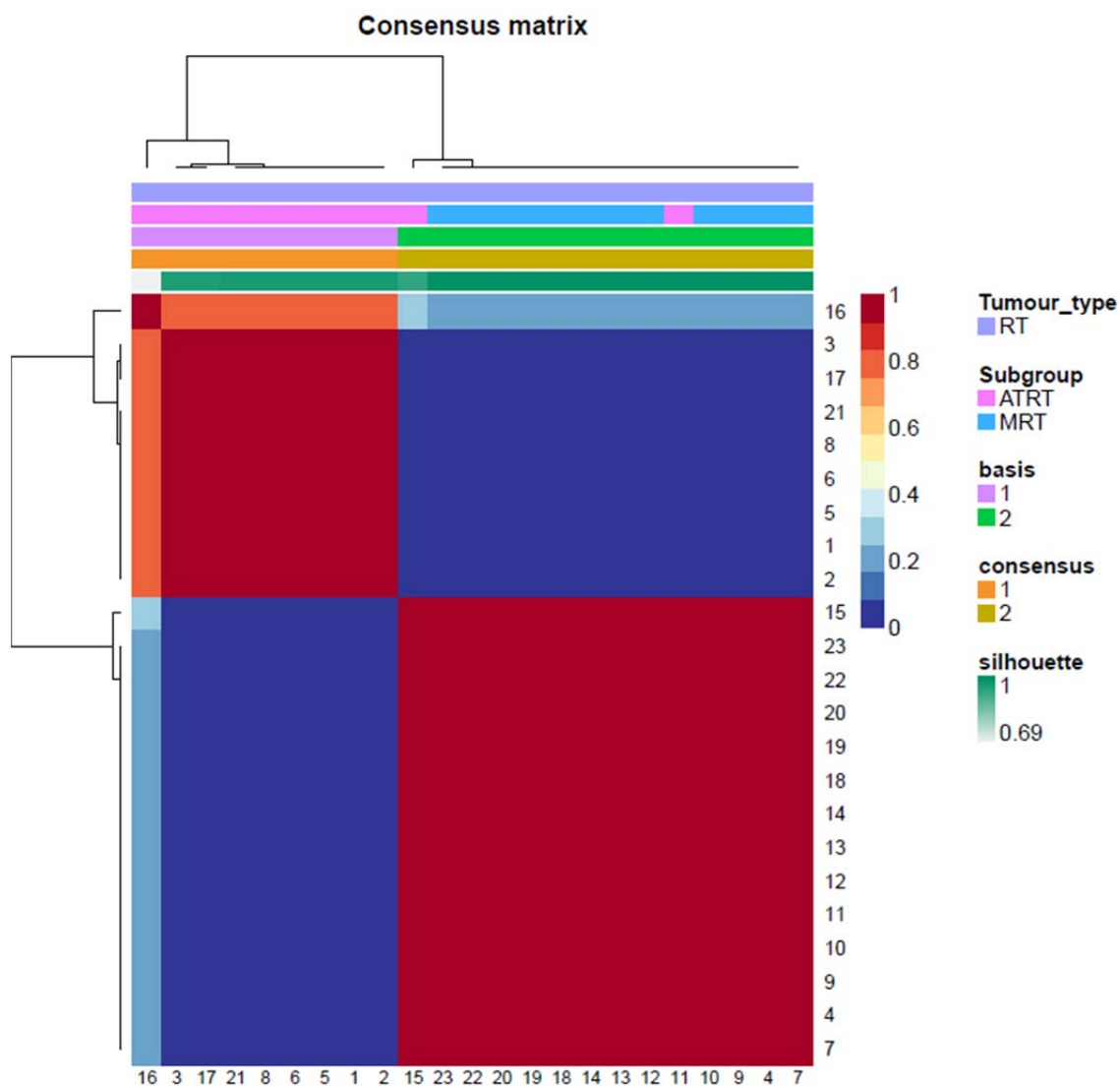


Figure 5.18 Unsupervised NMF and hierarchical cluster analysis ( $K=2$ ) on RNAseq data sets characterises two metagene classes mainly correlated with the location of the primary Rhabdoid Tumours. Extra Cranial Rhabdoid (ECRT) and Atypical Teratoid Rhabdoid Tumours (ATRT) are separated into two different groups. The matrixes are coloured 0 (deep blue, indicating samples never in the same cluster) to 1 (dark red, indicating samples always in the same cluster).

The heatmap in Figure 5.19 illustrates the gene pattern between different Extra Cranial Rhabdoid Tumours and Atypical Teratoid Rhabdoid Tumours.

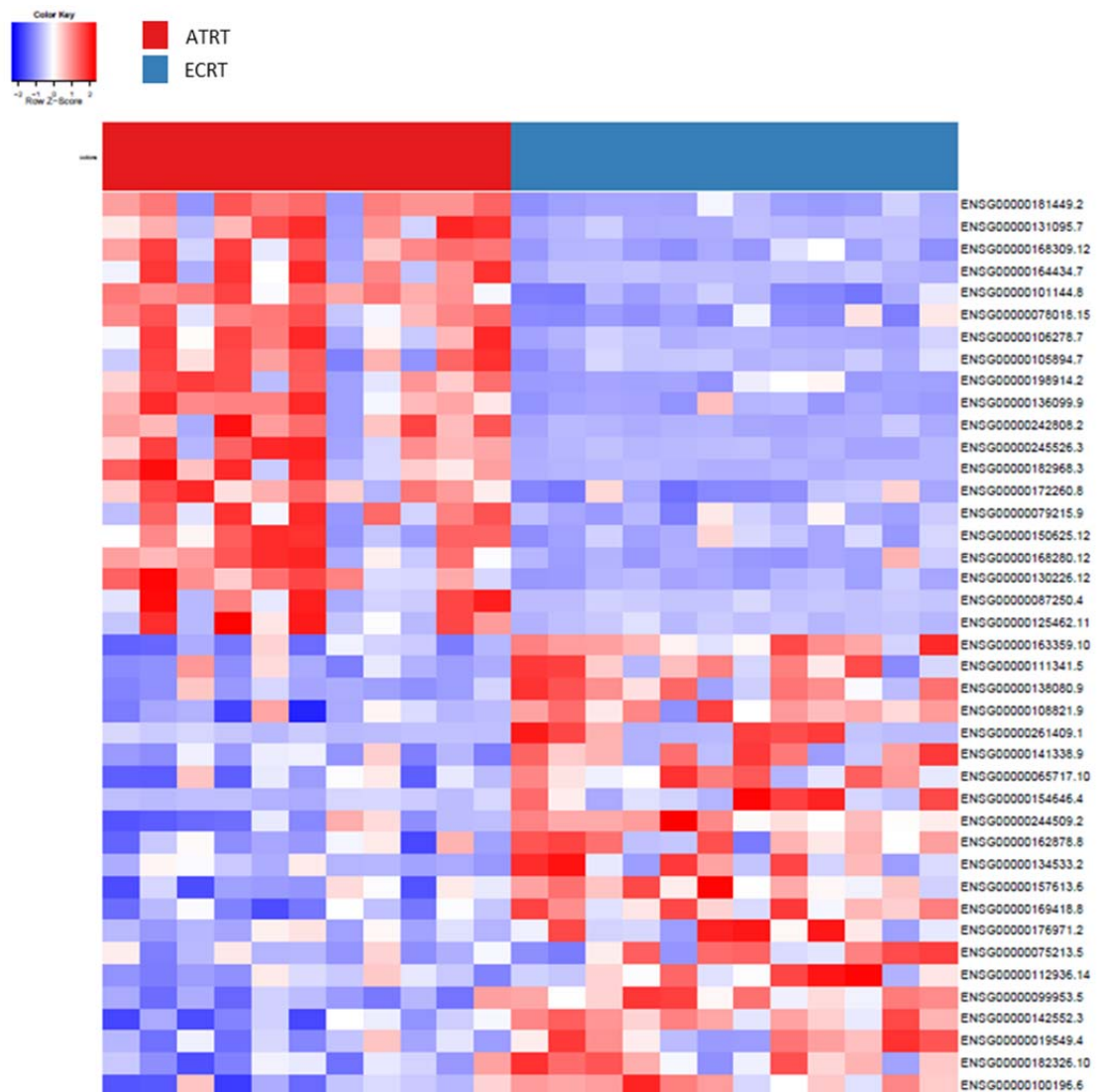
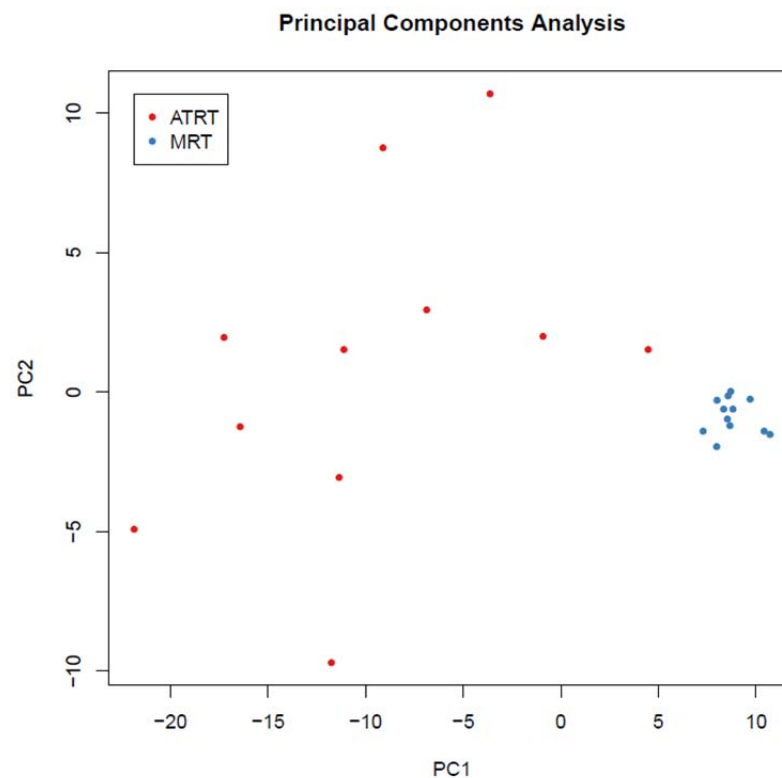


Figure 5.19 Heatmap of gene expression showing the 20 most highly expressed and differentially regulated genes between Extra Cranial Rhabdoid (ECRT) and Atypical Teratoid Rhabdoid Tumours (ATRT) of the brain. Each row represents a gene and each column a sample. Overexpressed (pale red to dark red) and under-expressed transcripts (pale blue to dark blue).

### *Supervised analysis*

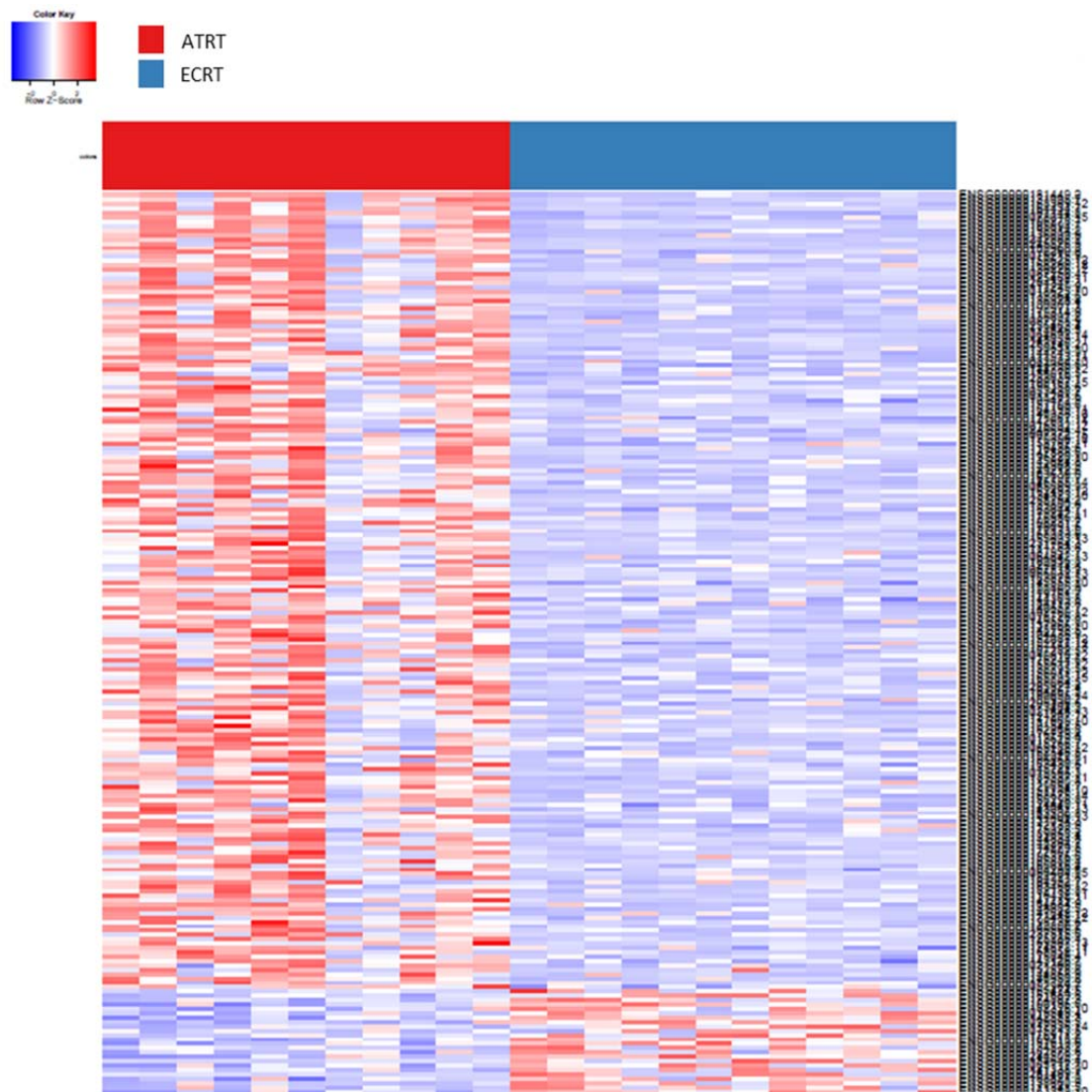
DEseq results were filtered down to 243 significantly differentially expressed genes (adjusted  $p$ -value  $< 0.01$ ; moderated log fold change  $\leq -2$  and  $\geq +2$ ). PCA analysis confirmed the presence of two subgroups within the rhabdoid tumours (Figure 5.20).



*Figure 5.20 Supervised principal component analysis on DEseq data set identifies two major subgroups in Rhabdoid cohort. In red are represented Atypical Teratoid Rhabdoid Tumours (ATRT) primary sample, while in blue the Extra Cranial Rhabdoid Tumours (ECRT). The x-axis is the first component; the y-axis is the second principal component.*



Filtered data were used to generate a supervised heatmap (Figure 5.21). Supervised analysis demonstrated that Extra Cranial Rhabdoid and Atypical Teratoid Rhabdoid Tumours have different gene patterns and the presence in the cohort of few Atypical Teratoid Rhabdoid Tumours with mix expression features.



*Figure 5.21 Heatmap of gene expression showing the highly expressed and differentially regulated genes between Extra Cranial Rhabdoid (ECRT) and Atypical Teratoid Rhabdoid Tumours (ATRT) of the brain. Each row represents a gene and each column a sample. Overexpressed (pale red to dark red) and underexpressed transcripts (pale blue to dark blue).*

DEseq results were filtered for values with an adjusted  $p$ -value greater than 0.0001 and moderated log fold change less than -4 or greater than 4. In Table 5.2 are indicated the 20 most differentially expressed genes after filtering and ranked by relative difference in expression. These genes are indicated in Figure 5.22 and in Figure 5.23. Eleven were associated with tumour progression (\*), two involved in cell proliferation (+), three directly involved in tumour invasion and angiogenesis (-), one indicated as tumour suppressor genes ( $\pm$ ), one gene indicated as oncogene and four acting in neurodevelopment proliferation of neural progenitor cells (#). Complete list in Appendix I.

Gene Symbol	Adjusted <i>p</i> value	Moderated log2fold
<b>TMPRSS15*</b>	3.81E-11	6.104431
<b>GJB2*-</b>	2.41E-07	4.489234
<b>RERG*</b>	1.83E-06	4.172613
<b>COL6A3 *+</b>	3.94E-08	4.139891
<b>FIBIN *-</b>	1.29E-05	4.103162
<b>EMILIN1</b>	1.19E-07	3.865497
<b>ABCA8</b>	2.68E-05	3.558311
<b>C7</b>	6.48E-05	3.42491
<b>MGP</b>	0.000452	3.374596
<b>SEMA3A* -</b>	0.000713	3.173271
<b>RFX4 *-</b>	2.45E-16	-7.12075
<b>NCAN</b>	3.87E-18	-7.19482
<b>PMP2*</b>	5.43E-16	-7.29577
<b>C1orf61~+</b>	1.85E-17	-7.32828
<b>MT3*-</b>	4.60E-17	-7.338
<b>SALL3#</b>	3.71E-15	-7.36159
<b>PTPRZ1</b>	1.09E-20	-7.62713
<b>GFAP*#</b>	6.94E-25	-8.42908
<b>FABP7*#</b>	2.76E-22	-8.64199
<b>SOX1 ± #</b>	6.23E-32	-9.83142

*Table 5.2 Differentially expressed gene in Extra-cranial Rhabdoid (ECRT) vs Atypical Teratoid Rhabdoid Tumours (ATRT). Moderated log2fold change is given as a ratio between ECRT and ATRT. Functional annotation of genes relevant to cancer are shown in the table with \* for genes associated to tumor progression ,+ involved in cell proliferation,- directly correlated with tumour invasion and angiogenesis, ~ oncogene, # for genes acting in neurodevelopment proliferation of neural progenitor cells, ± tumor suppressor genes.*

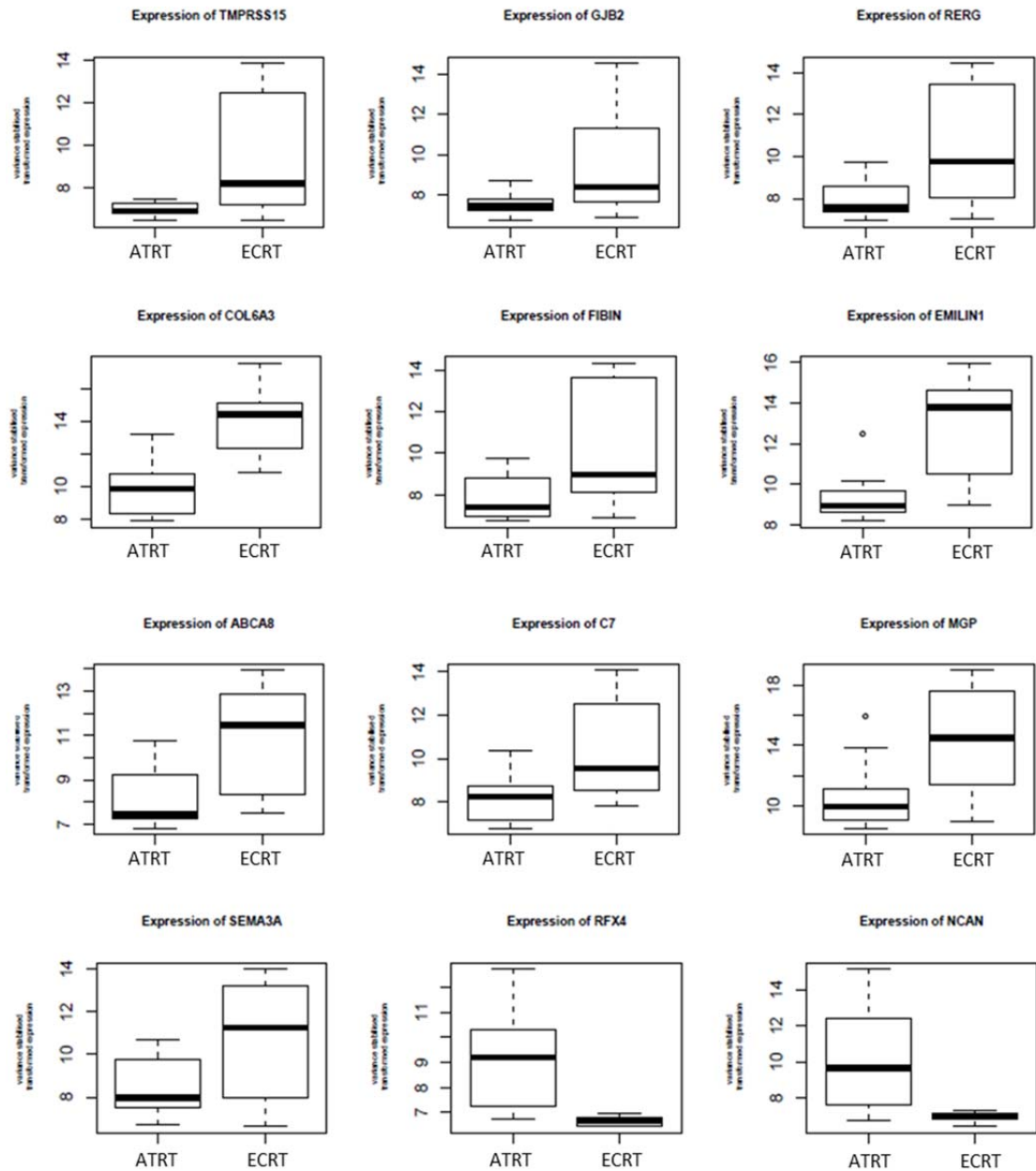


Figure 5.22 Relative differences in expression of 1-12 most deregulated genes between Extra Cranial Rhabdoid Tumours (ECRT) and Atypical Teratoid Rhabdoid Tumours (ATRT). Boxplots show Variant stabilised transformed expression across the samples. Boxplots show median score (thick black line), range (whiskers) and inter-quartile ranges (extent of box).

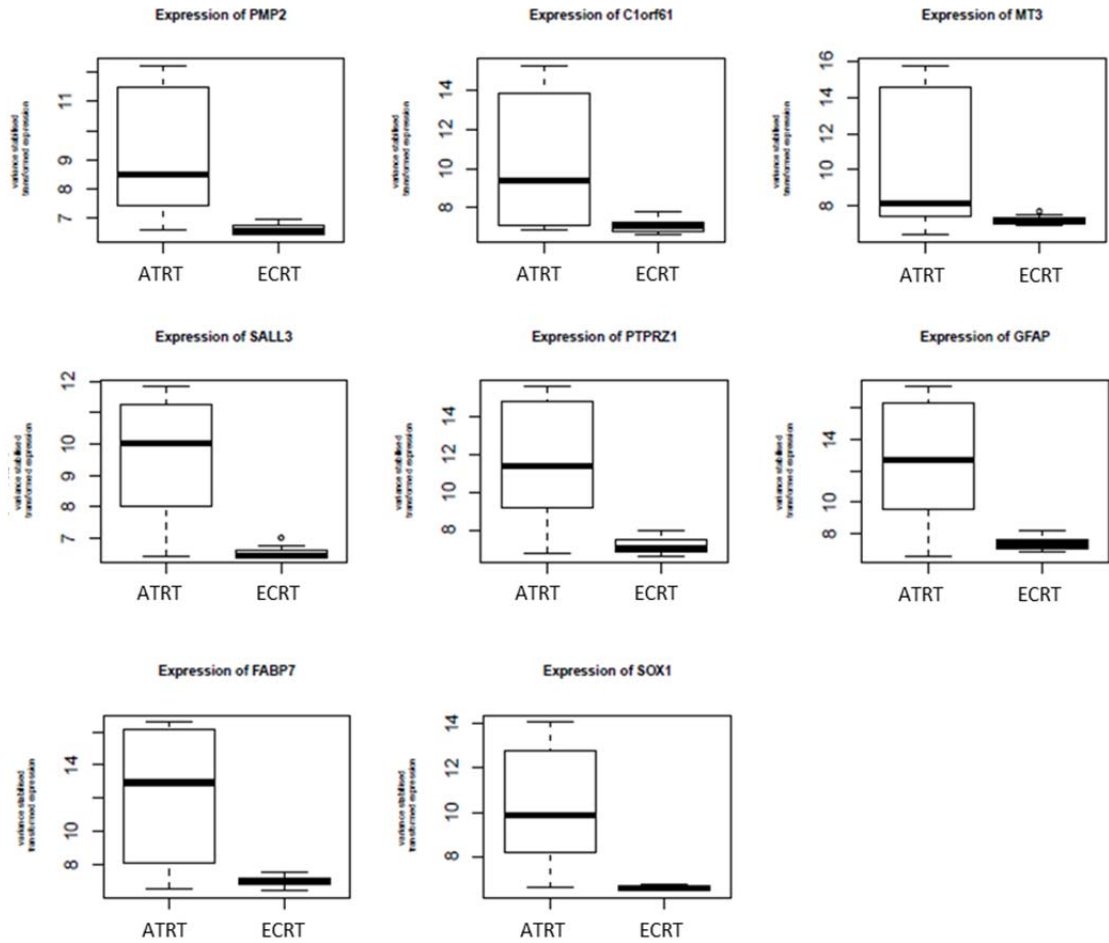


Figure 5.23 Relative difference in expression of 13-20 most deregulated genes between Extra Cranial Rhabdoid Tumours (ECRT) and Atypical Teratoid Rhabdoid Tumours (ATRT). Boxplots show Variant stabilised transformed expression across the samples. Boxplots show median score (thick black line), range (whiskers) and inter-quartile ranges (extent of box).

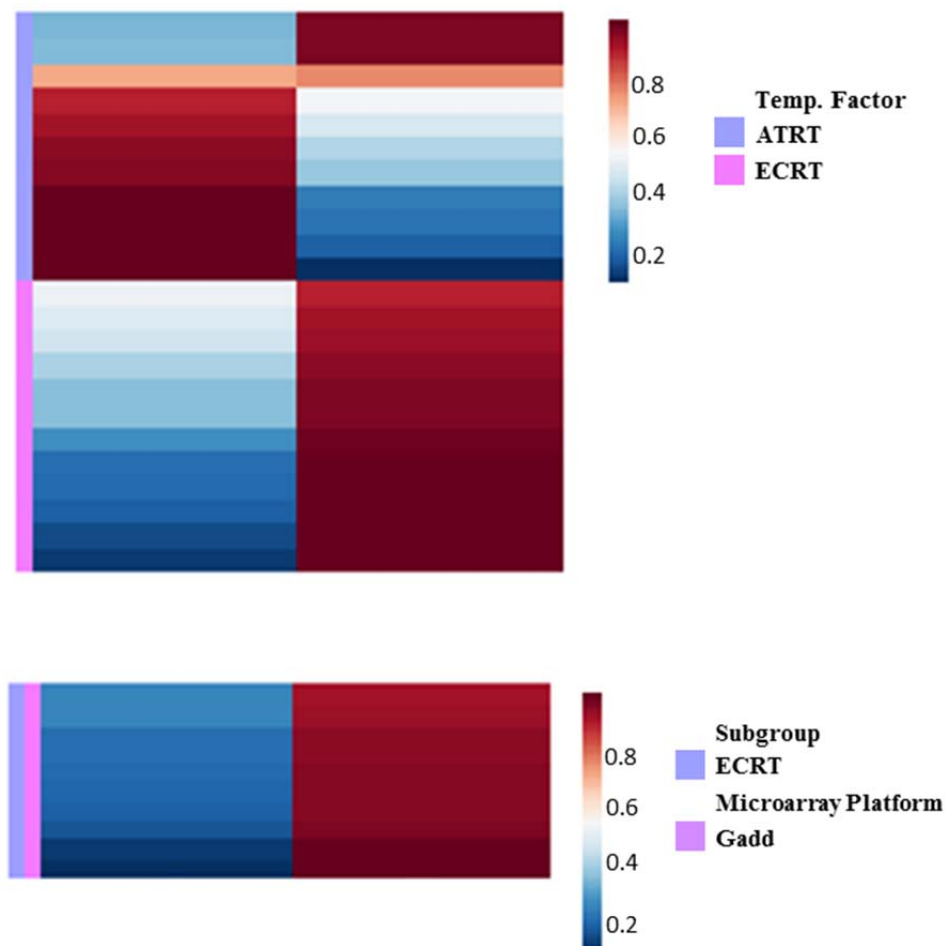
Interestingly Atypical Teratoid Rhabdoid Tumours showed up-regulation of those genes related with neurodevelopment, proliferation of neural progenitor cells, and neuro-oncogenesis. For instance, *GFAP* is expressed in Glioblastoma and Chordomas (Theurillat, Hainfellner et al. 1999, Santos, Carvalho et al. 2009), moreover its up regulation is correlated with increased expression of the *B-FABP* family in Glioma (Godbout, Bisgrove et al. 1998). SOX family proteins have a key role in embryonic and postnatal development and lead tumourigenesis in a wide number of cancers; in hepatocellular carcinoma hypermethylation of *SOX1* causes abnormal activation of Wnt/ $\beta$ -catenin signalling pathway and is a marker for cortical neural progenitor cells (Elkouris, Balaskas et al. 2011, Tsao, Yan et al. 2012).

#### *RNA-seq metagene projection onto existing primary Malignant Rhabdoid Tumours microarray data sets*

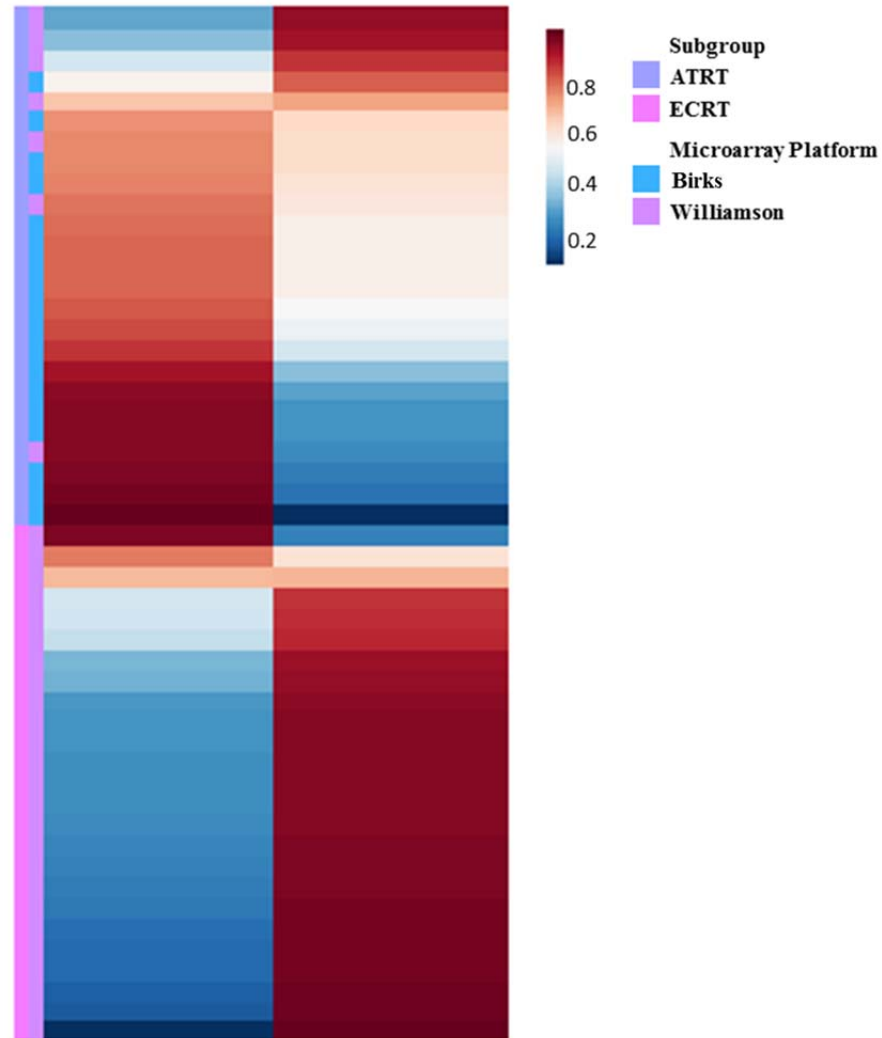
The metagene difference between Extra Cranial Rhabdoid Tumours and Atypical Teratoid Rhabdoid Tumours was further studied by interactive exploration of gene expression measurements obtained by RNA-seq in combination with pre-existing gene expression analysis. In particular, the metagene patterns were projected onto three data set (Birks, Kleinschmidt-DeMasters et al. 2010, Gadd, Sredni et al. 2010, Williamson unpublished data), using a data-driven learning process.

The three data sets came from different laboratories and were acquired on different microarray technologies, Affymetrix HGU133plus2 for Birks and Williamson's data set and Affymetrix HGU133 for the Gadd's cohort. Metagene projection methodology was applied as previously described by Tamayo et al. (Tamayo, Scanfeld et al. 2007). For the analysis a rank of  $k = 2$  as metagene factors was used, in accordance with my previous analysis of cophenetic correlation and dispersion measurement in my RNA-seq data set. Fig. 2 shows the resulting heat maps for the test sets and projected model. This analysis showed clearly the two metagene associate with the Extra Cranial and the brain biological phenotypes across technology platform although not exclusively. In particular Extra Cranial Rhabdoid Tumours constitute mainly a unique and clear subgroup,

although some Extra Cranial Rhabdoid and Atypical Teratoid Rhabdoid Tumours samples present features of both metagenes. This confirms the difference in gene signature we observed in our cohort.



*Figure 5.24 Heatmaps of metagene projection of Malignant Rhabdoid Tumour primary samples. These heatmaps show the metagene expression levels for each sample. Extra Cranial Rhabdoids (ECRT) are indicated in pink while the Atypical Teratoid Rhabdoid Tumours (ATRT) are indicated in purple. Heat maps of the RNA-seq projected model (top) and of the Gadd's Cohort after metagene projection (bottom) are indicated. The test data was acquired using HGU133a microarray platforms on a total of 10 ECRT.*



*Figure 5.25 Heatmap of metagene projection of Malignant Rhabdoid Tumour primary samples. These heat maps show the metagene expression levels for each sample. Extra Cranial Rhabdoids (ECRT) are indicated in pink while the Atypical Teratoid Rhabdoid Tumours (ATRT) are indicated in Purple. Heat maps of Birks' and Williamson's data set after metagene projection is indicated. Both data tests were acquired using HGU133plus2 microarray platforms. In blue is indicated Birks's cohort, constituting a total of 18 ECRT. Williamson data set is indicated in dark pink and it is composed of 23 Extra Cranial primary tumours and 7 Atypical Teratoid Rhabdoid Tumours.*



### 5.3.5 DEXSeq analysis: differential exon usage

DEXSeq divides reads mapped to the same gene into contigs depending on the exons overlapped. Each count does not coincide with exons but with exonic parts since transcripts often present exons with different sizes in the same genomic location. Therefore two transcripts characterised by different lengths for the same exon are assigned two counts (one for the shortest exon, and one for the longer one), plus the counts for the remaining exons.

DEXseq, as well as DEseq, computes False Discovery Rate (FDR), fold change and adjusted p-values which are subjected to the Benjamini-Hochberg correction for multiple testing. The data set was firstly tested for variance dispersion, p-value distribution and FDR, as previously described (section 5.3.3). Figure 5.26 represents the estimate of dispersion in our data set, where the black dots represent the empirical dispersion values and the red line the fitted values. The means across the genes were equally distributed, as indicated by the dispersion trend. A differential exon usage significance test between two conditions (Rhabdoid vs Medulloblastoma primary tumours and Extra Cranial Rhabdoid Tumours vs Atypical Teratoid Rhabdoid Tumours) was calculated (Figure 5.26, Figure 5.27, Figure 5.28, Figure 5.29).

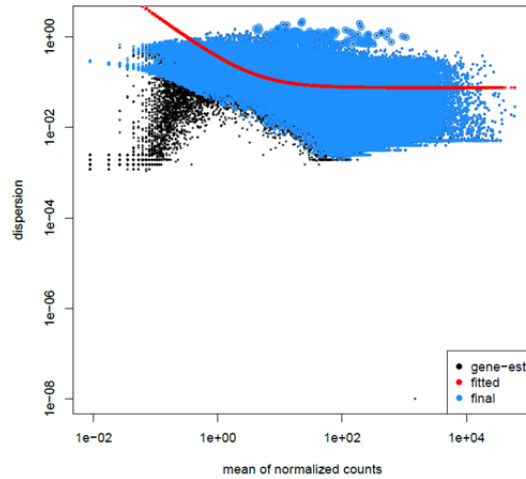


Figure 5.26 Differential exon usage dispersion estimation and model fitting of Rhabdoid and Medulloblastoma derived DEXseq data. Each dot represents exonic part dispersion estimation and the red line the fitted mean dispersion function.

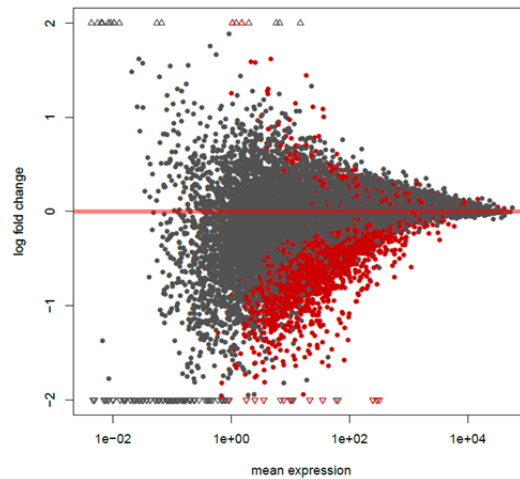


Figure 5.27 Per exonic part dispersion estimates of Rhabdoid and Medulloblastoma derived DEXseq data. Differentially used exons Log2fold change versus mean expression over all samples. Red dots represent exonic parts significantly differentially used in 10% FDR in Rhabdoid primary tumours, when the Benjamini and Hochberg method is applied. The lines on the top and on the bottom of the plot represent exon usage with a very low or very high log fold change (the value off scale).

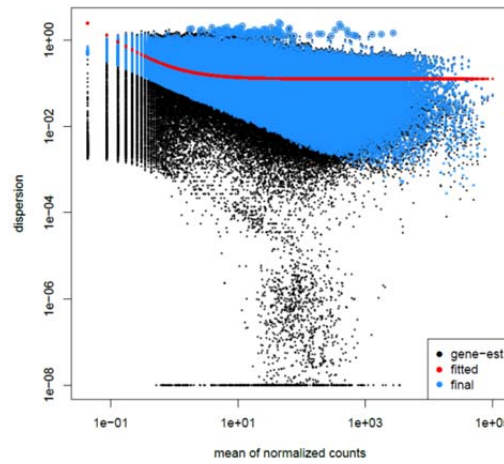


Figure 5.28 Differential exon usage dispersion estimation and model fitting of Rhabdoid Tumour DEXseq data. Each dot represents exonic part dispersion estimation and the red line the fitted mean dispersion function.

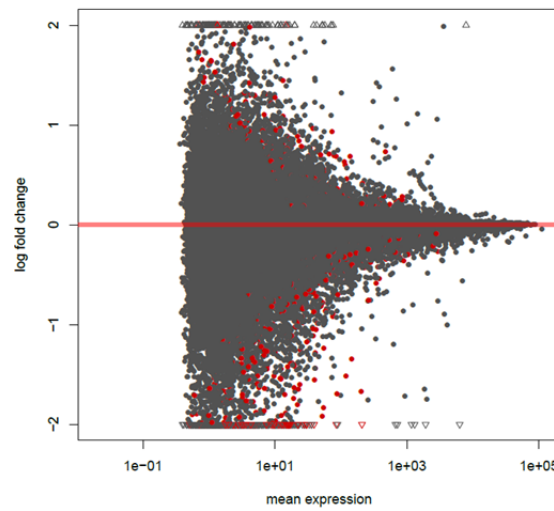


Figure 5.29 Differentially used exons Log2fold change versus mean expression over all samples. Red dots represent exonic parts differentially used at 10% FDR, when Benjamini and Hochberg method is applied. The lines on the top and on the bottom of the plot represent genes with a very low or very high log fold change (the value off scale).

#### *5.3.5.1 Detection of differentially used exons between Malignant Rhabdoid Tumours and Medulloblastoma primary tumours*

The DEXseq analysis provides a list of genes and their respective exonic parts, described by adjusted  $p$ -value (calculated using Benjamini and Hochberg method) and log2 fold change (calculated by dividing the mean of Rhabdoid primary samples by the mean of Medulloblastoma primary samples.)

The DEXseq results were filtered for values with an adjusted  $p$ -value greater than 0.0001 and log2 fold change less than -2 or greater than 2, moreover only protein coding genes were analysed, resulting in 2002 significant genes with differential exon usage. In Table 5.3 are indicated the 30 most differentially used exons as a result of filtering (Full results can be found in Appendix 2). From the 30 genes in the table, four were associated with tumour progression, two with cell cycle, one indicated as potent tumour suppressor genes and one involved in autophagy in breast cancer and generally high expressed in brain tissue (<http://www.genecards.org/>). Complete list in Appendix I.

Gene Symbol	Exon	Adjusted <i>p</i> value	log2fold
<b>OSTM1</b>	E020	7.42E-05	2.691206
<b>GNE +</b>	E023	7.22E-07	2.446738
<b>ARGLU1 *</b>	E010	1.86E-07	2.001177
<b>BAHD1</b>	E001	1.00E-16	1.617837
<b>STK32C</b>	E016	6.43E-09	1.59003
<b>ACTR3B</b>	E009	8.79E-14	1.582486
<b>TUBA1C *</b>	E002	1.35E-06	-1.700879104
<b>SPRYD3</b>	E013	1.02E-08	-1.728031529
<b>FARSB</b>	E017	7.45E-15	-1.744469767
<b>MGST2</b>	E016	1.22E-09	-1.756456049
<b>HES6 ~</b>	E024	2.86E-06	-1.78873522
<b>DERL1</b>	E014	5.91E-07	-1.813588525
<b>C6orf89 +</b>	E006	9.62E-07	-1.819608749
<b>SYPL1</b>	E009	1.36E-83	-1.944165165
<b>OCIAD2 *</b>	E001	1.88E-06	-1.945484655
<b>LYRM2</b>	E007	7.74E-06	-1.951875084
<b>RPS27L #</b>	E001	1.61E-12	-1.958972254
<b>GOLIM4</b>	E012	3.79E-07	-2.012249694
<b>DNAJC1</b>	E006	2.27E-05	-2.024899986
<b>JOSD1</b>	E023	1.31E-06	-2.048115239
<b>SLC29A4</b>	E001	5.01E-05	-2.059570348
<b>MAP1LC3A *</b>	E002	7.07E-05	-2.13733565
<b>PPIA</b>	E022	1.68E-08	-2.261613246
<b>DHRS7B</b>	E001	1.71E-11	-2.265327266
<b>PPIA</b>	E021	3.83E-06	-2.312175449
<b>NCS1</b>	E002	3.18E-05	-2.318299372
<b>LDLRAD3</b>	E010	2.97E-05	-2.583835253
<b>MRPL32</b>	E010	1.72E-08	-2.963216251
<b>SLC48A1</b>	E004	1.69E-06	-9.939249198

*Table 5.3 Genes with differentially used exons in Rhabdoid vs Medulloblastoma primary tumours. Functional annotation of genes relevant to cancer is shown in the table with + for genes associated with cell cycle such as apoptosis, \* for genes associated with tumor progression, and # indicates genes related to autophagy in cancer and ~ tumour suppressor genes.*

*HES6* is an interesting candidate for further analysis since it has been associated with both genesis and progression of prostate cancer. In particular, it promotes growth and migration advantages to prostate cells by interaction with Notch signalling. Figure 5.30 shows the differential expression at the exon level of *HES6* obtained with DEXSeq. *HES6* in Rhabdoid primary samples shows in particular different usage at exon 1 (contigs 25-24 and 22-16) resulting in the usage of alternative transcripts start sites.

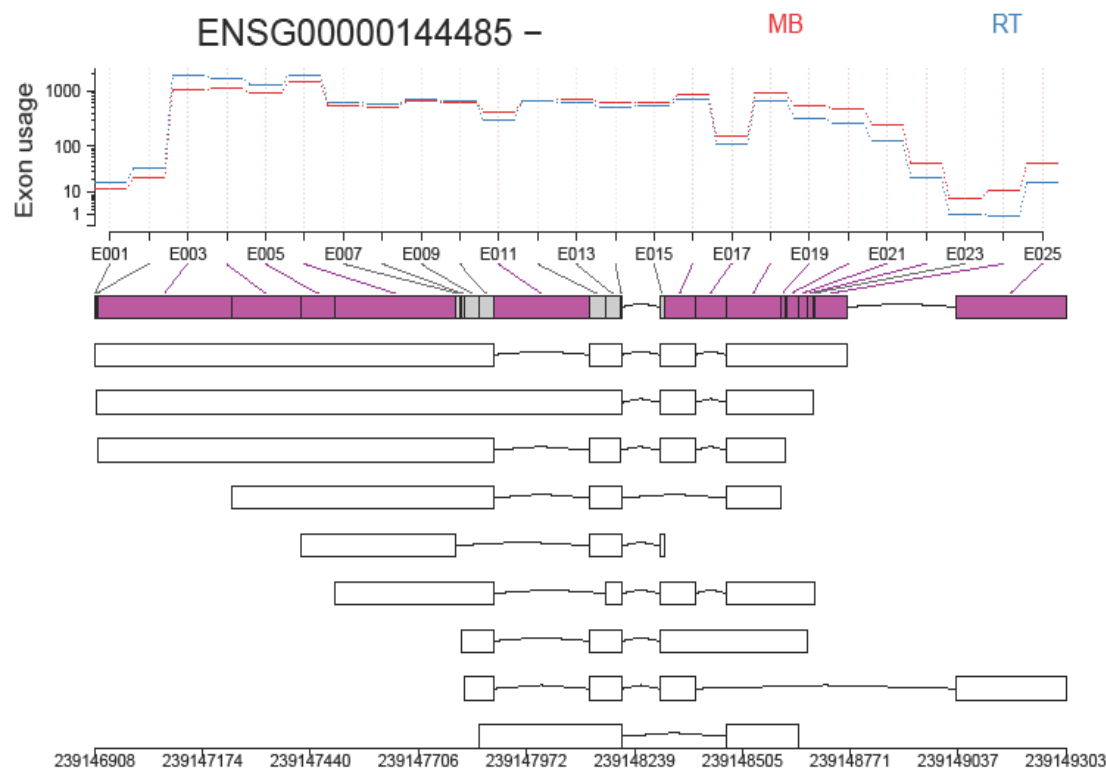
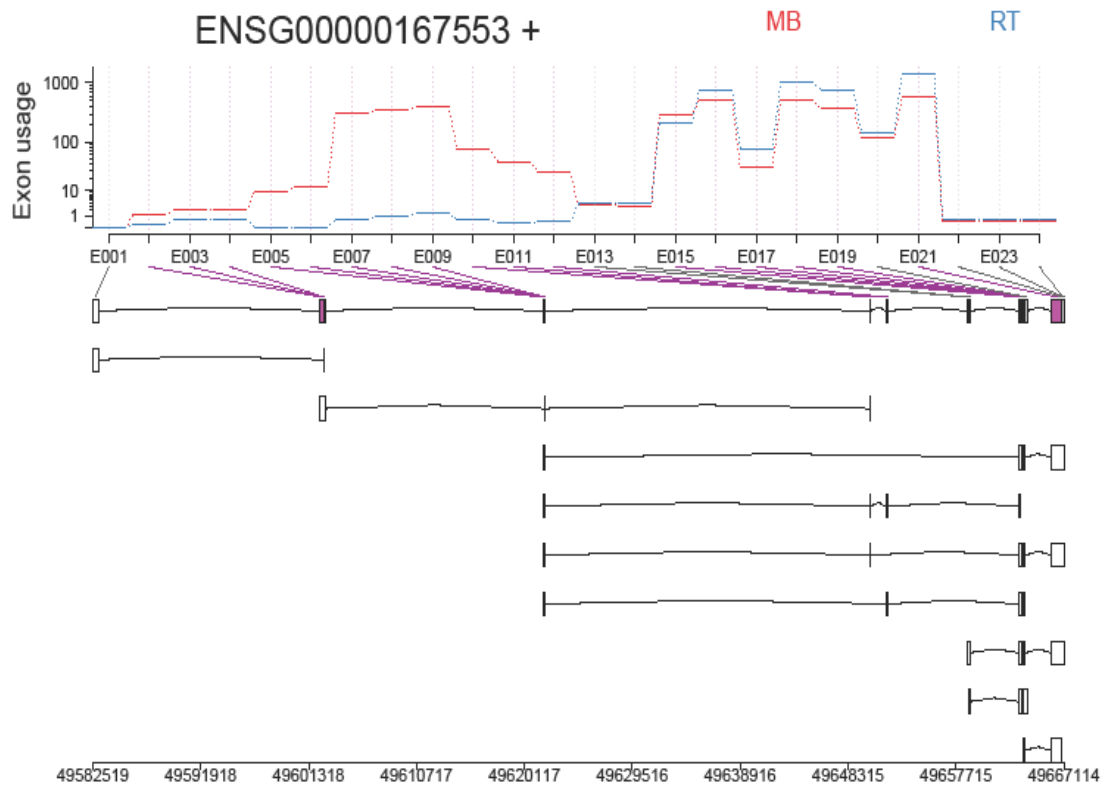


Figure 5.30 *HES6* shows differential exon usage in Rhabdoid Tumours compared with Medulloblastoma sample (exonic parts 16-22 and 24-25). The red and blue lines do not correspond directly to exons, but to exonic parts, Exon usage (read counts normalised to the overall expression of the gene) is indicated in red for Medulloblastoma patients and in blue for Rhabdoid primary tumours. At the bottom is included an annotated transcript model.

The *TUBA1C* gene was found to be highly expressed (Caracciolo, D'Agostino et al. 2010) and hypermethylated in Medulloblastoma cell lines (Dolores Hamilton, data unpublished). In Figure 5.31 Medulloblastoma primary sample show an increase usage especially of the Exonic Parts 5-12, this effect results in aberrant expression, while in the Rhabdoid samples exon usage is significantly elevated in exonic parts 16-19 and 21.



*Figure 5.31 TUBA1C shows differential exon usage in Medulloblastoma tumours compared with Rhabdoid sample (exonic parts 5-12). The red and blue lines do not correspond directly to exons, but to exonic parts, Exon usage (read counts normalised to the overall expression of the gene) is showed in red for Medulloblastoma patients and in blue for Rhabdoid primary tumours. At the bottom is included an annotated transcript model*

The *RPS27L* gene is a nuclear protein which is found up-regulated in a variety of cancers, including breast, colon and liver cancer (Ganger, Hamilton et al. 1997, Ganger, Hamilton et al. 2001, Atsuta, Aoki et al. 2002). However *RPS27L* also has a tumour suppressor activity, since it regulates *p53* activity by abrogating Mdm2-induced *p53* degradation (He and Sun 2006). Figure 5.32 shows the differential exon usage of *RPS27L* gene in Rhabdoid Tumours (Exonic parts 1,5,6,10,11,15); noticeably the first transcript contains an open reading frame.

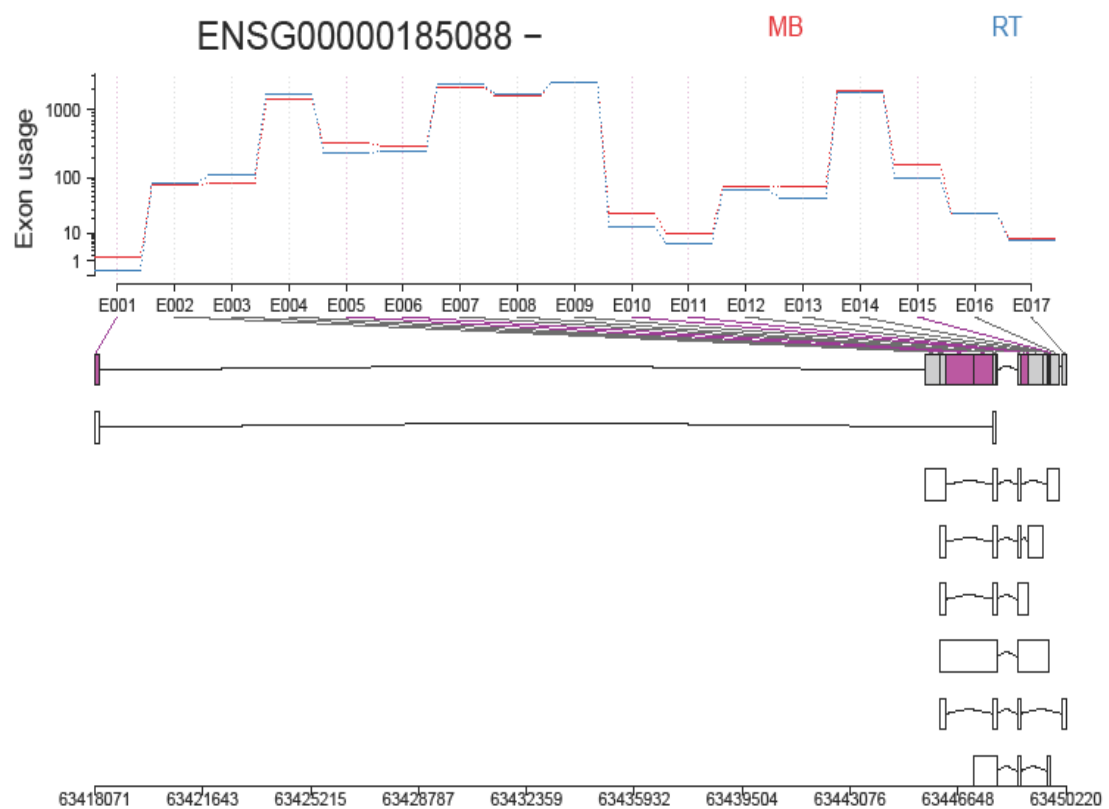


Figure 5.32 Differential exon usage of *RPS27L* in Rhabdoid sample tumours (Exonic parts 1,5, 6, 10, and 11). The red and blue lines do not correspond directly to exons, but to exonic parts, Exon usage (read counts normalised to the overall expression of the gene) is indicated in red for Medulloblastoma patients and in blue for Rhabdoid primary tumours. At the bottom is included an annotated transcript model.



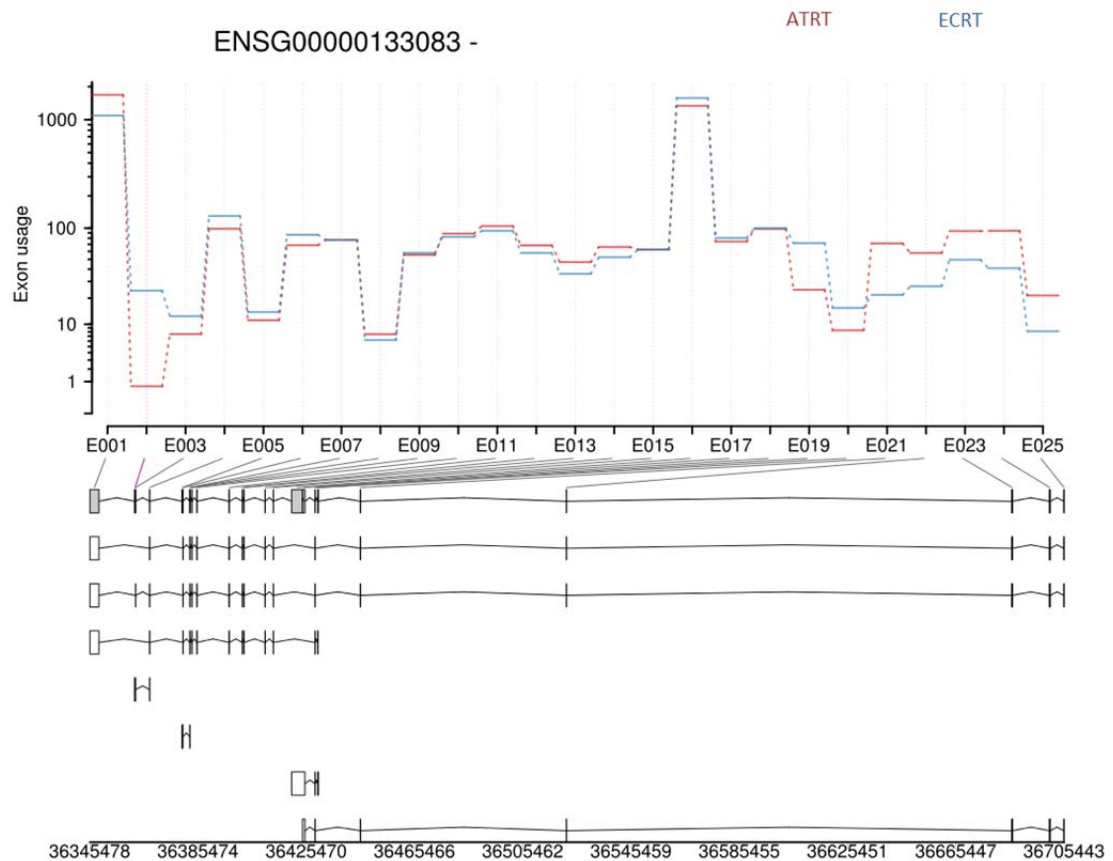
#### 5.3.5.2 *Detection of differentially used exons between Extra Cranial Rhabdoids tumours and Atypical Teratoid Rhabdoid Tumours.*

The DEXseq data set was filtered for values with an adjusted *p*-value greater than 0.01 and log2 fold change less than -2 or greater than 2, moreover only protein coding gene were analysed. This filtering characterises 561 genes with significant differential exon usage. In Table 5.4 are indicated the 30 most differentially used exons as result of filtering. From the 30 genes in the table, five were associated with tumour progression , two gene involved in Wnt pathway , one associated with regulation of pluripotency and angiogenetic factors and one directly regulated by MicroRNA.

*DCLK1* and *RNF220* are two interesting candidates: both are involved in Notch and Wnt pathway regulation (Chandrakesan, Weygant et al. 2014, Ma, Yang et al. 2014, Qu, May et al. 2014). Interestingly, the relationship between these two genes has not been studied, however this analysis shows in Extra Cranial Rhabdoid Tumours an increase usage in exonic part 2 affecting exon number 2 (Figure 5.33), while *RNF220* usage is decreased from exonic part 1 to 17 affecting exons number 1, 4 , 5 and 9 (Figure 5.34) . This indicated a possible effect on the Wnt pathway and /or Notch signalling regulation via differential exon usage. Complete list in Appendix I.

Gene Symbol	Exon	Adjusted <i>p</i> value	log2fold
<b>MATR3#</b>	E015	0.000620502	9.541355489
<b>FRMD4B</b>	E038	0.003767638	2.686008583
<b>LIMCH1+</b>	E015	0.003066572	2.46342767
<b>ABR</b>	E099	0.000639142	2.214450991
<b>DCLK1~</b>	E002	1.44E-05	2.173069156
<b>MBP*</b>	E043	0.000620502	2.015855785
<b>AC024560.3</b>	E001	0.019082575	-5.796040636
<b>CCT7</b>	E001	8.84E-05	-5.80737613
<b>CCT7</b>	E002	4.06E-05	-5.944955916
<b>RFTN1</b>	E032	0.002379	-6.528080996
<b>SPAG16</b>	E038	1.04E-05	-6.991812374
<b>SPAG16</b>	E039	1.41E-06	-7.08877491
<b>SUMF1</b>	E010	9.93E-05	-7.471278383
<b>PRCP</b>	E017	9.55E-06	-7.521447266
<b>SUMF1</b>	E012	3.10E-05	-7.579783724
<b>RCN1*</b>	E008	6.34E-07	-7.77565421
<b>RNF220 +</b>	E003	1.79E-05	-10.43948157
<b>KAT7</b>	E019	2.44E-05	-10.49322879
<b>UBA6-AS1</b>	E028	6.91E-07	-10.50189736
<b>SEMA4A *</b>	E001	3.41E-05	-15.91598894
<b>MLLT4*</b>	E017	0.000153	-17.31556728
<b>HOPX*</b>	E025	7.07E-05	-6.99181

*Table 5.4 Genes with differentially used exon in Rhabdoid primary tumours*  
*Functional annotation of genes relevant to cancer are shown in the table*  
*with + for genes involved in Wnt pathway, \* for genes associated with tumor*  
*progression, and # indicates genes regulated by MicroRNA and ~gene*  
*regulating pluripotency and angiogenetic factors genes.*



*Figure 5.33 Differential exon usage of DCLK1 in Rhabdoid sample tumours (Exonic part 2). The red and blue lines do not correspond directly to exons, but to exonic parts, Exon usage (read counts normalised to the overall expression of the gene) is indicated in red for Atypical Teratoid Rhabdoid Tumours (ATRT) and in blue for the Extra Cranial Rhabdoid Tumours (ECRT). At the bottom is included an annotated transcript model.*

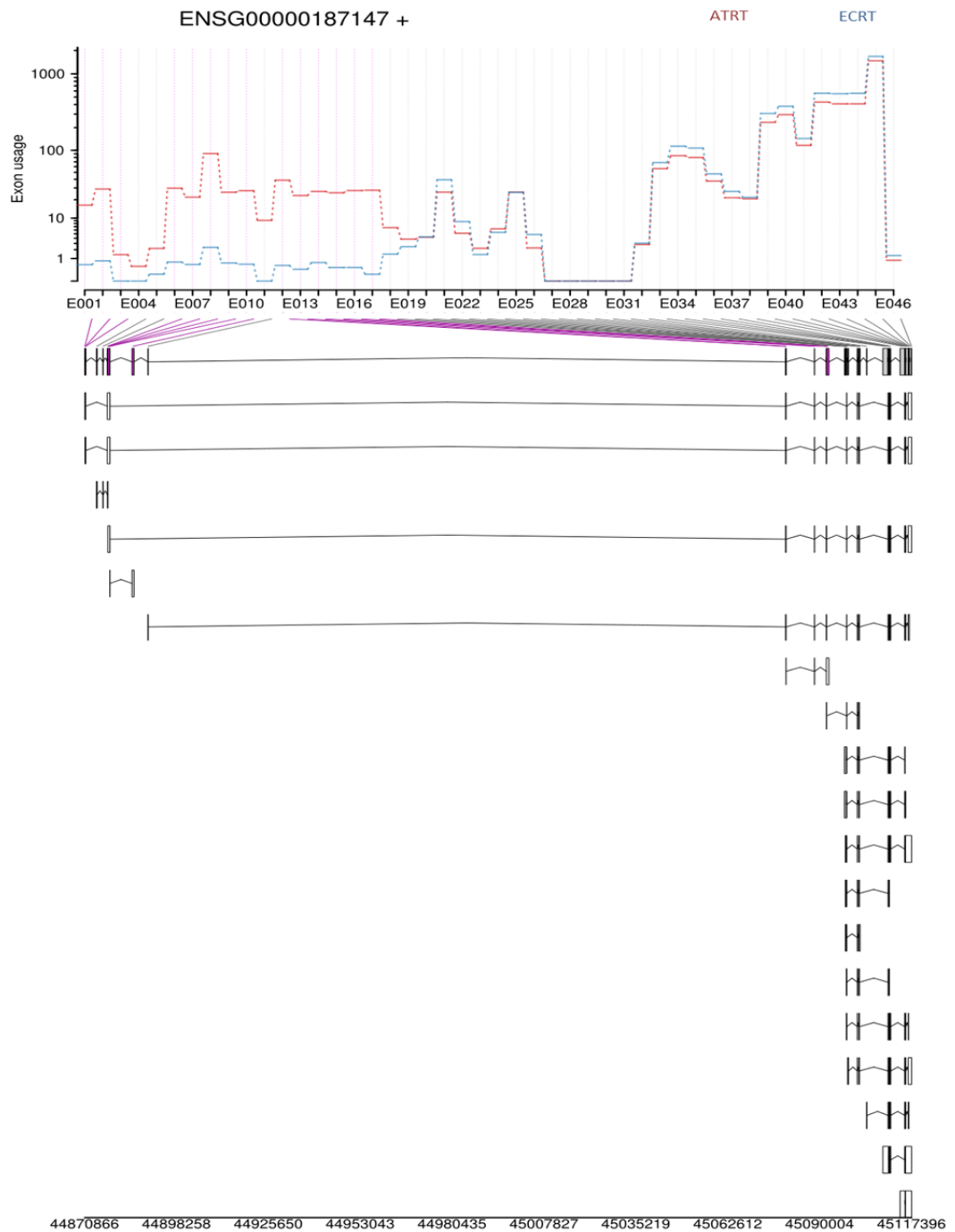
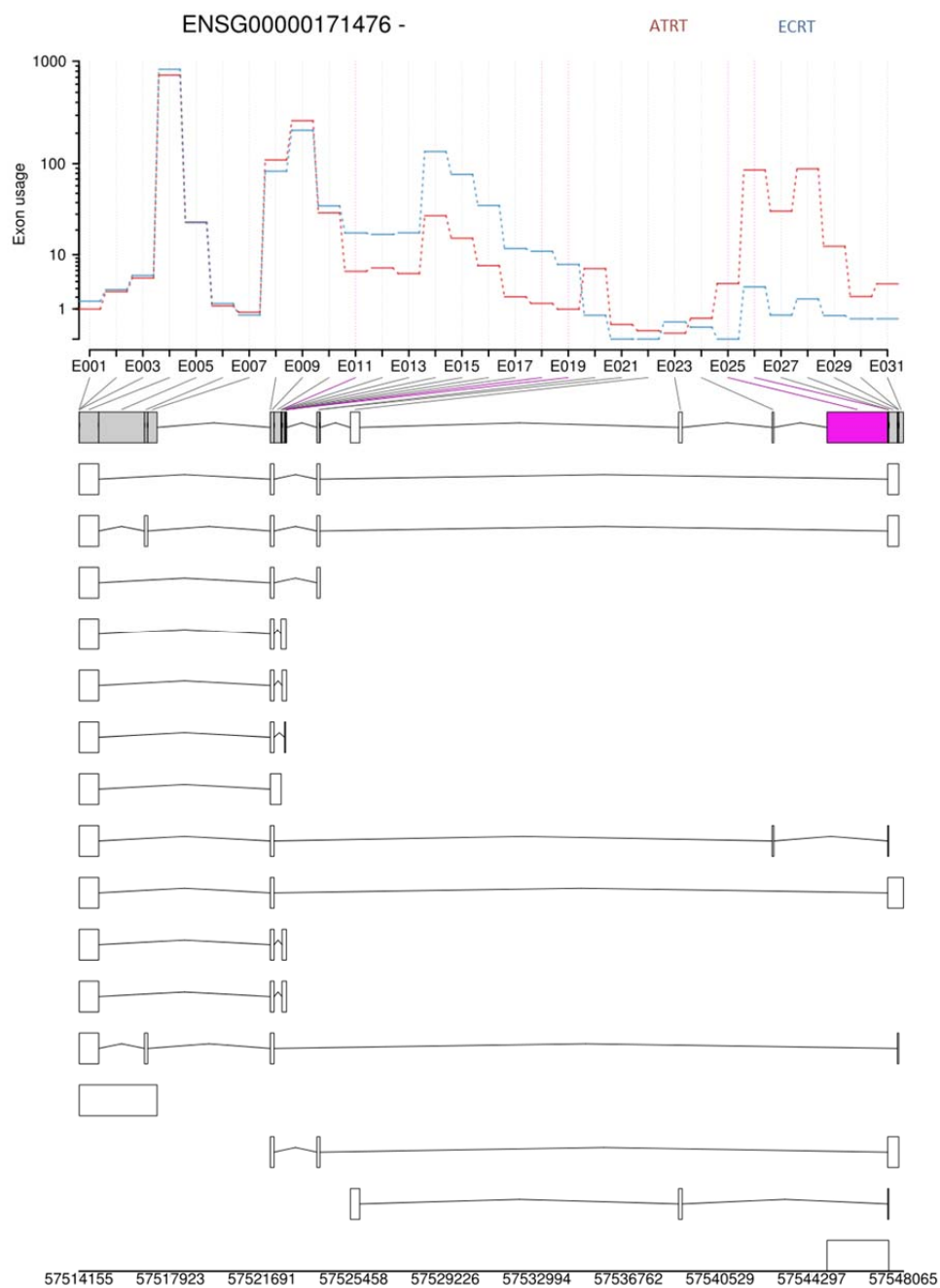


Figure 5.34 Differential exon usage of RNF220 in Rhabdoid sample tumours (exonic parts 1- 17). The red and blue lines do not correspond directly to exons, but to exonic parts, Exon usage (read counts normalised to the overall expression of the gene) is indicated in red for Atypical Teratoid Rhabdoid Tumours (ATRT) and in blue for the Extra Cranial Rhabdoid Tumours (ECRT). At the bottom is included an annotated transcript model.

*HOPX* has been indicated as tumour suppressor gene in a variety of cancer, such as pancreatic, colorectal and lung cancer (Waraya, Yamashita et al. 2012, Cheung, Zhao et al. 2013, Yamashita, Katoh et al. 2013) and found methylated in poor differentiated colorectal, oesophageal and gastric cancer (Yamashita, Kim et al. 2008, Ooki, Yamashita et al. 2010, Katoh, Yamashita et al. 2012). *HOPX* is a critical nodes of Wnt pathway in lung tumours (Pacheco-Pinedo, Durham et al. 2011). Figure 5.35 shows the differential expression at the exon level of *HOPX* obtained from DEXSeq analysis. *HOPX* in Extra Cranial Rhabdoid Tumours samples shows in particular significant differential exon usage at Exonic parts 18-19 and 25-26, to determine differential recruitment of an alternative splicing.



*Figure 5.35 Differential exon usage of HOPX in Rhabdoid Tumours (Exonic parts 11- 17; 25 and 26). The red and blue lines do not correspond directly to exons, but to exonic parts, Exon usage (read counts normalised to the overall expression of the gene) is indicated in red for Atypical Teratoid Rhabdoid Tumours (ATRT) and in blue for the Extra Cranial Rhabdoid Tumours (ECRT). At the bottom is included an annotated transcript model.*

## 5.4 450K methylation array confirm different biological group in CNS tumours

This section aimed to test the hypothesis that differential patterns of methylation exist within Rhabdoid Tumours. Characterisation of methylation pattern can improve the understanding of the epigenetic mechanism involved in *SMARCB1* loss. A total of 39 arrays were analysed on 450K methylation microarray (25 Atypical Teratoid Rhabdoid Tumours, 14 in Extra Cranial Rhabdoid Tumours).

### 5.4.1 *Primary samples preparation and sequencing*

43 Rhabdoid primary samples were collected from multiple centres throughout the UK in collaboration with the CCLG bio bank ([www.cclg.org.uk/tissue-bank](http://www.cclg.org.uk/tissue-bank)). All samples were evaluated at the referring centre and subsequently centrally review by CCLG bio bank. DNA was extracted as per described methods (1.1.1).

### 5.4.2 QC analysis

The quality of the data was analysed using the Minfi package, at a pre-processing stage. First, bisulfite conversion was evaluated: incomplete or failure of the bisulfite conversion might results in biases to the analysis. This process consists of evaluation of the red and green control channel raw data, converting them into methylated and unmethylated signal. Quality analysis is computed separately for the InfI (Infinium I) and InfII (Infinium II) control probes (**Error! Reference source not found.** ).

$\beta$  values were derived using the BASH algorithm (section 4.2.2.3): Beta values can be fully unmethylated (0), fully methylated (1). In order to filter anomalous methylation intensities, the distribution of beta values were further analysed to determine the quality of the data. Overall, the majority of samples from our cohort proved to be of good

quality with less than one percent of CpG sites with detection  $p$ -values (indicating good signal from an individual priobe) equal or greater than 0.001 (Figure 5.37)

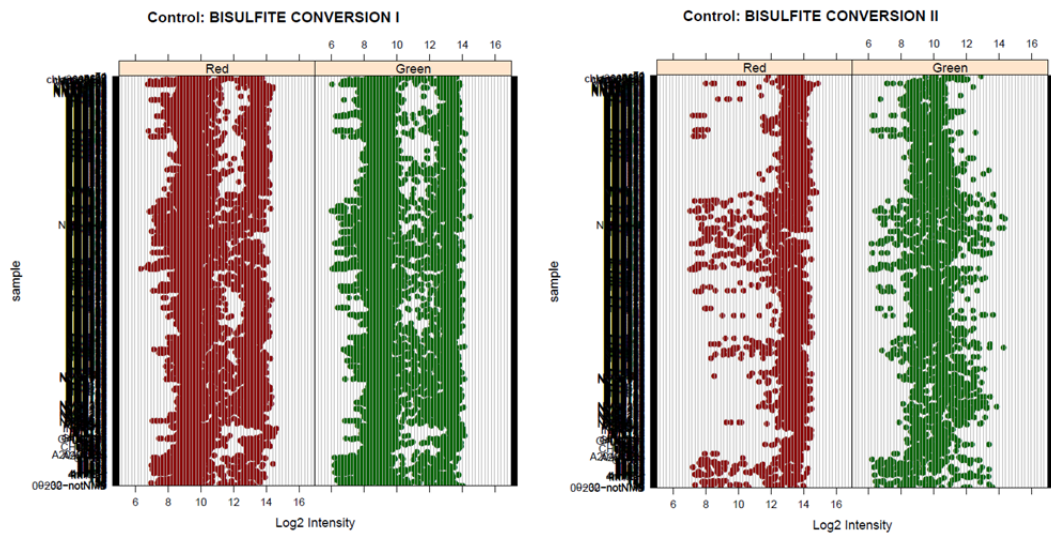
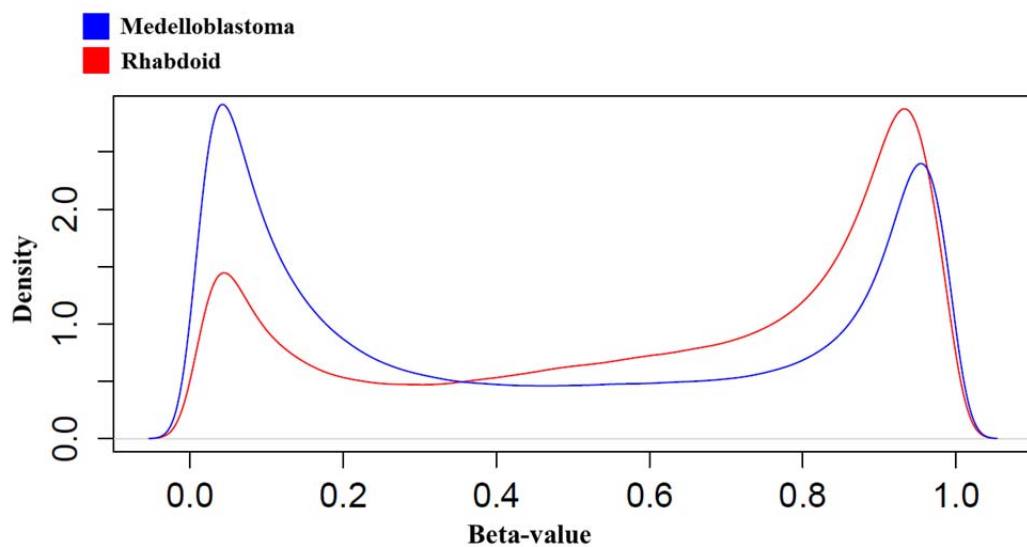


Figure 5.36 Control strip plot of 450K methylation data. Quality evaluation for *InfI* (left) and *InfII* (right) is indicated. Distribution of signal intensity for each of the quality control probes in both is plotted. Control probes in successfully converted samples present similar distribution in both channels.





*Figure 5.37 Overall density distributions of beta values and the density distributions in primary Rhabdoid cohort. Kernel density plot illustrates the bimodal distribution of  $\beta$ -values for primary tumours analysed.*

### 5.4.3 Global patterns of DNA methylation in Rhabdoid Tumours

The cohort set, 26290/40000 probes ( $p < 0.01$ ) showed evidence of methylation ( $\beta$  score  $> 0.333$ ) in one or more samples with a mean beta value of 0.59. The distribution of average methylation status within test cohorts is summarised in Table 5.5. In particular, there was a very marked difference in the methylation status of probes located within and outside of CpG islands: only 5311/40000 probes were located within CpG islands, in contrast the majority of the probes were located outside the CpG Island.

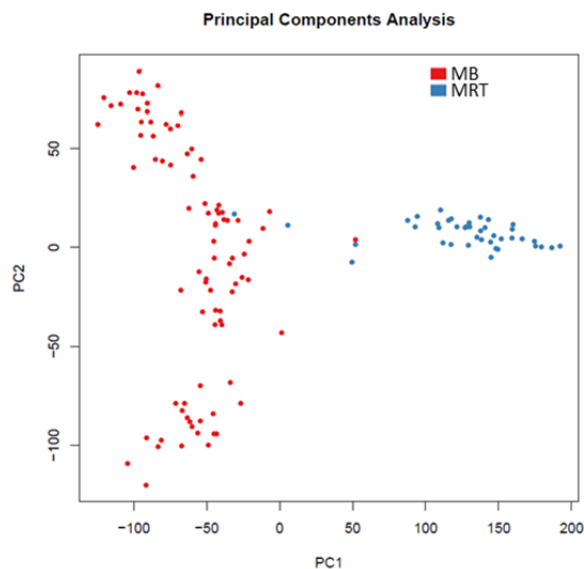
	Probes classification	%
<b>All probes</b>	Methylated	7029 (11.6%)
	Hemimethylated	12599 (57%)
	Unmethylated	5762 (31.5%)
<b>CpG Island</b>	Methylated	1659 (6.3%)
	Hemimethylated	2561 (9.7%)
	Unmethylated	1091 (4.15%)
<b>Non CpG Island</b>	Methylated	5370 (20.4%)
	Hemimethylated	10038 (38.0%)
	Unmethylated	4671 (17.7%)

*Table 5.5 Global methylation distribution among Malignant Rhabdoid primary tumours and Medulloblastomas test cohorts (26290 probes). Average methylation scores were calculated for all probes across the training and test cohort. For three probe categories (all probes, probes located within CpG islands and probes located outside of CpG islands), the number of unmethylated (average  $\beta < 0.333$ ), methylated (average  $\beta > 0.667$ ) and hemi-methylated probes ( $0.333 < \text{average } \beta \leq 0.666$ ) are shown for the total cohort.*

## 5.4.4 Cluster analysis identifies 5 distinct DNA methylation subgroups

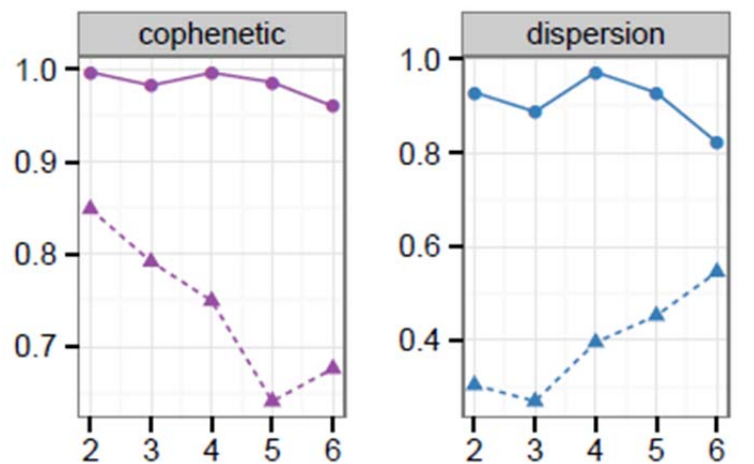
### 5.4.4.1 Unsupervised analysis

In order to explore differences in methylation status between Rhabdoid and Medulloblastoma primary tumours an unsupervised clustering approach (PCA, HC, and NMF) was used. The rationale behind choosing Medulloblastoma as a comparator group was the same as for the expression studies, namely it was a related embryonal tumour type. Principal component analysis (PCA) was performed on a subset of probes selected as the most variable CpG probes within the data set. The PCA plot shows the tendency of that data to cluster into two biological groups: a group formed by Rhabdoid samples and one mainly constituted by Medulloblastoma (Figure 5.38).



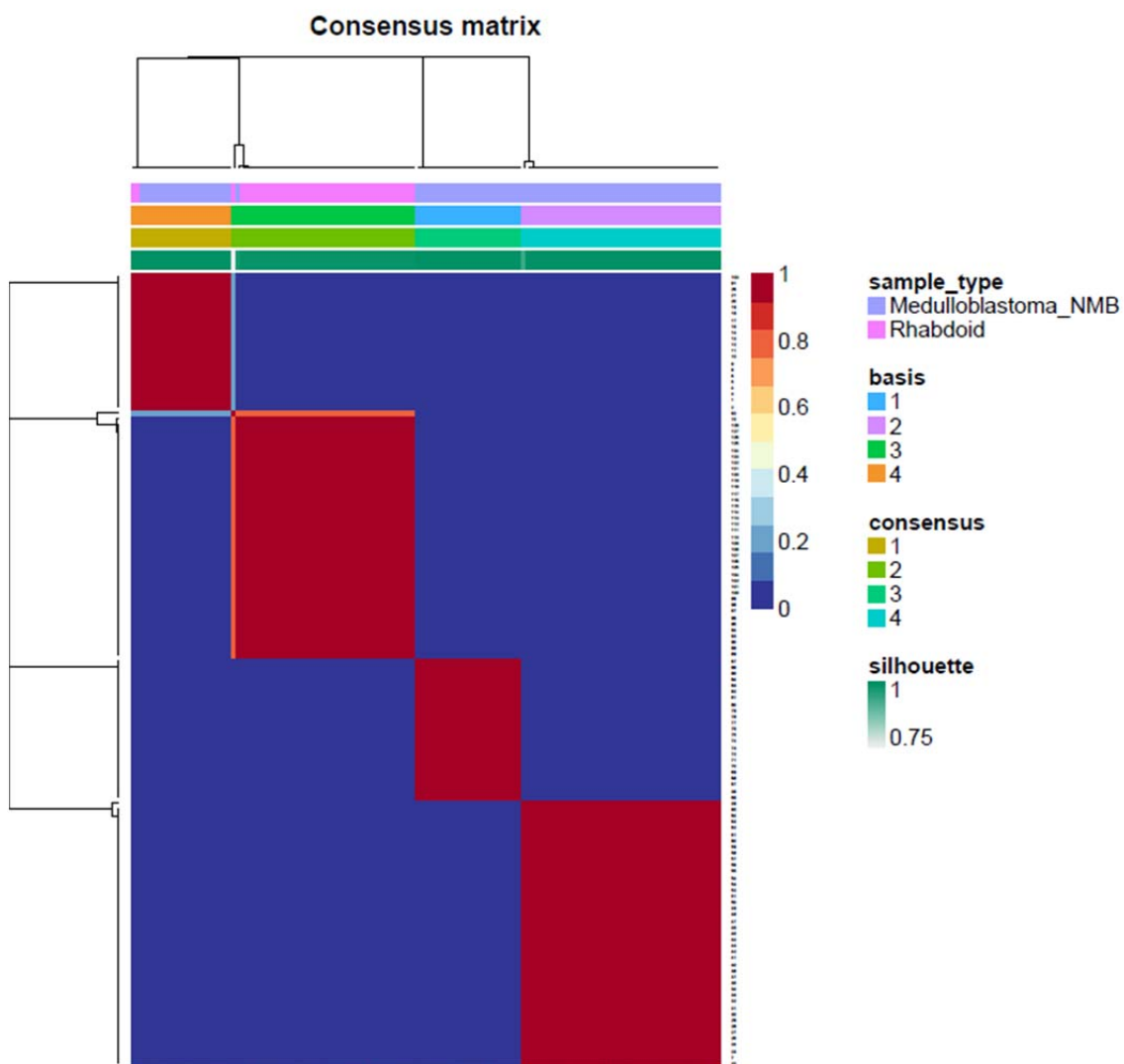
*Figure 5.38 Unsupervised principal component analysis of 450K methylation profiles of primary Medulloblastoma and Rhabdoid Tumours distinguishes two biological groups. Points in red represent Medulloblastoma primary samples while Rhabdoid Tumours are indicated in blue. The x-axis is the first component; the y-axis is the second principal component.*

A consensus clustering approach (repetitions =100) was Initially applied to the data set to identify the optimal number of metagenes and clusters. The stability of the cluster was determined by evaluating cophenetic correlation and dispersion coefficient, as previously described. Metagenes were computed between ranks (number of metagenes) 2 to 6 on the most variable CpG probes. The most robust consensus clustering was chosen as k=4 (Figure 5.39).



*Figure 5.39 Cophenetic correlation and dispersion measures to assess the stability of clustering associated with each rank k for the methylation data set. Dotted line shows randomised while data undotted shows actual data.*

Consensus NMF and unsupervised hierarchical clustering confirmed the pattern observed in the expression data: that Rhabdoid Tumours mainly (41/43) form a unique group regardless of the location of the malignancy, with the exception of two samples (Figure 5.40). These two sample are unde review to asses if they are misdiagnosed Atypical Teratoid Rhabdoid Tumours or they present low expression of *SMARCB1*.



*Figure 5.40 Unsupervised NMF and hierarchical cluster analysis (K=4) on Rhabdoid and Medulloblastoma derived 450K methylation data set. Rhabdoid and Medulloblastoma are clustered in different groups, the genetic complexity of Medulloblastoma (characterized by 4 different subgroups: WNT, SHH, Group 3 and Group 4) are retained in this analysis. The matrices were coloured 0 (deep blue, indicating samples never in the same cluster) to 1 (dark red, indicating samples always in the same cluster).*

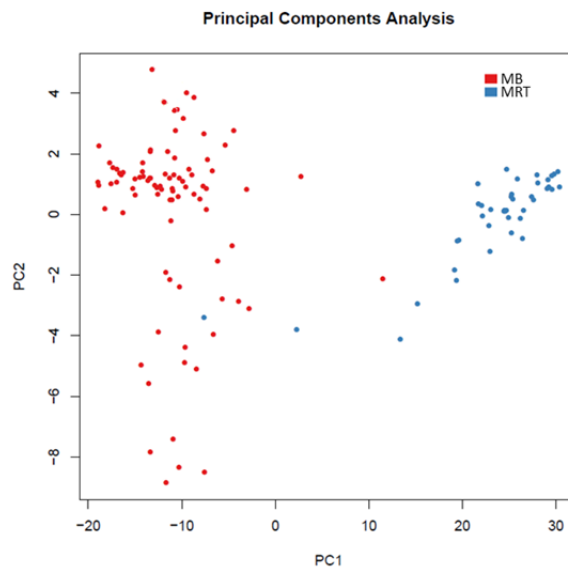
The heatmap indicated in Figure 5.41 contains the top 100 differentially methylated probes (as defined by delta value) and illustrates the difference in methylation between Medulloblastoma and Rhabdoid primary tumours.



*Figure 5.41 Heatmap of methylation showing the 100 most differentially methylated probes between Medulloblastoma and Rhabdoid Tumours. Each row represents a CpG probe and each column a sample. Hypermethylated (pale red to dark red) and Hypomethylated CpGs (pale blue to dark blue). Medulloblastoma primary samples are represented by side bars in Red while Rhabdoid patients are indicated in blue.*

#### 5.4.4.2 Supervised analysis

To further investigate the difference between Medulloblastoma and Rhabdoid Tumours methylation data were filtered in order to obtain a subset composed only of those CpGs which were most significantly differentially methylated. PCA analysis was computed using a subset of 3267 probes. These confirmed the magnitude of the difference in methylation between Rhabdoid and Medulloblastoma primary tumours (Figure 5.42) and also the two Medulloblastoma samples under investigation remained outliers from the group.



*Figure 5.42 Supervised principal component analysis of 450K methylation profiles of primary Medulloblastoma and Rhabdoid Tumours distinguishes two biological groups. Points in red represent Medulloblastoma primary samples while Rhabdoid Tumours are indicated in blue. The x-axis is the first component; the y-axis is the second principal component.*

In order to highlight the most Rhabdoid specific methylation events. The results were filtered for CpG probes with an adjusted  $p$ -value greater than 0.0001 and a mean delta (mean difference) beta value less than -0.6 or greater than 0.6. In Table 5.6 are indicated the 20 most differentially methylated probes as result of filtering and their associated genes, relative difference in methylation status are indicated in Figure 5.43 and Figure 5.44 (complete list at appendix 2). Five were associated with tumour progression (\*), three with cell cycle (+), one indicated as candidate tumour suppressor genes (~) and two involved in neuro-development (#) (<http://www.genecards.org/>). Complete list in Appendix I.

Probes	Gene Symbol	Average Delta values	LogFold change	Adjusted <i>p</i> value
<b>cg02899346</b>	RTN2	0.744696	7.413367	9.38E-39
<b>cg19378416</b>	OGFOD3	0.72519	6.34529	8.72E-43
<b>cg25820257</b>	RTN2	0.717721	6.302051	1.74E-43
<b>cg00092518</b>	PHF11~	0.711333	6.985098	1.43E-33
<b>cg01063813</b>	STAT6*	0.7081	6.464976	8.35E-41
<b>cg01005968</b>	RTN2	0.628437	4.981123	6.86E-36
<b>cg10253465</b>	RTN2	0.626719	5.184367	2.67E-38
<b>cg05367028</b>	PRKAG2*	-0.60039	-4.589	2.79E-41
<b>cg14792075</b>	KIF26B*	-0.60148	-4.53975	1.22E-31
<b>cg13036593</b>	TNXB*	-0.60257	-4.99831	7.79E-32
<b>cg26758486</b>	SDK2	-0.75876	-8.27617	6.78E-33
<b>cg20181739</b>	SH3RF3	-0.7637	-7.03077	2.94E-47
<b>cg04389398</b>	ORC5	-0.76494	-7.23292	3.08E-42
<b>cg14992000</b>	COL20A1 +	-0.76503	-7.22481	1.67E-51
<b>cg15084433</b>	SND1+	-0.76827	-7.49857	3.37E-47
<b>cg13524328</b>	FBXW8*+ #	-0.78451	-6.90548	2.09E-52
<b>cg01807313</b>	FBXW8	-0.79331	-7.49232	3.64E-49
<b>cg16111737</b>	FBXW8	-0.79347	-7.24504	1.71E-51
<b>cg27396830</b>	DRAXIN#	-0.79429	-7.70371	1.67E-51
<b>cg25620901</b>	HEATR2	-0.81378	-8.26217	4.13E-48

*Table 5.6 Selected significantly differentially methylated probes in Rhabdoid vs Medulloblastoma primary tumours (adjust *p*value <0.01; Mean Delta values  $\geq +0.6$   $\leq -0.6$ ). Functional annotation of genes relevant to cancer is shown in the table with + for genes associated with cell cycle such as apoptosis, \* for genes associated with tumour progression, and # indicates genes related to autophagy in cancer and ~ tumour suppressor genes.*



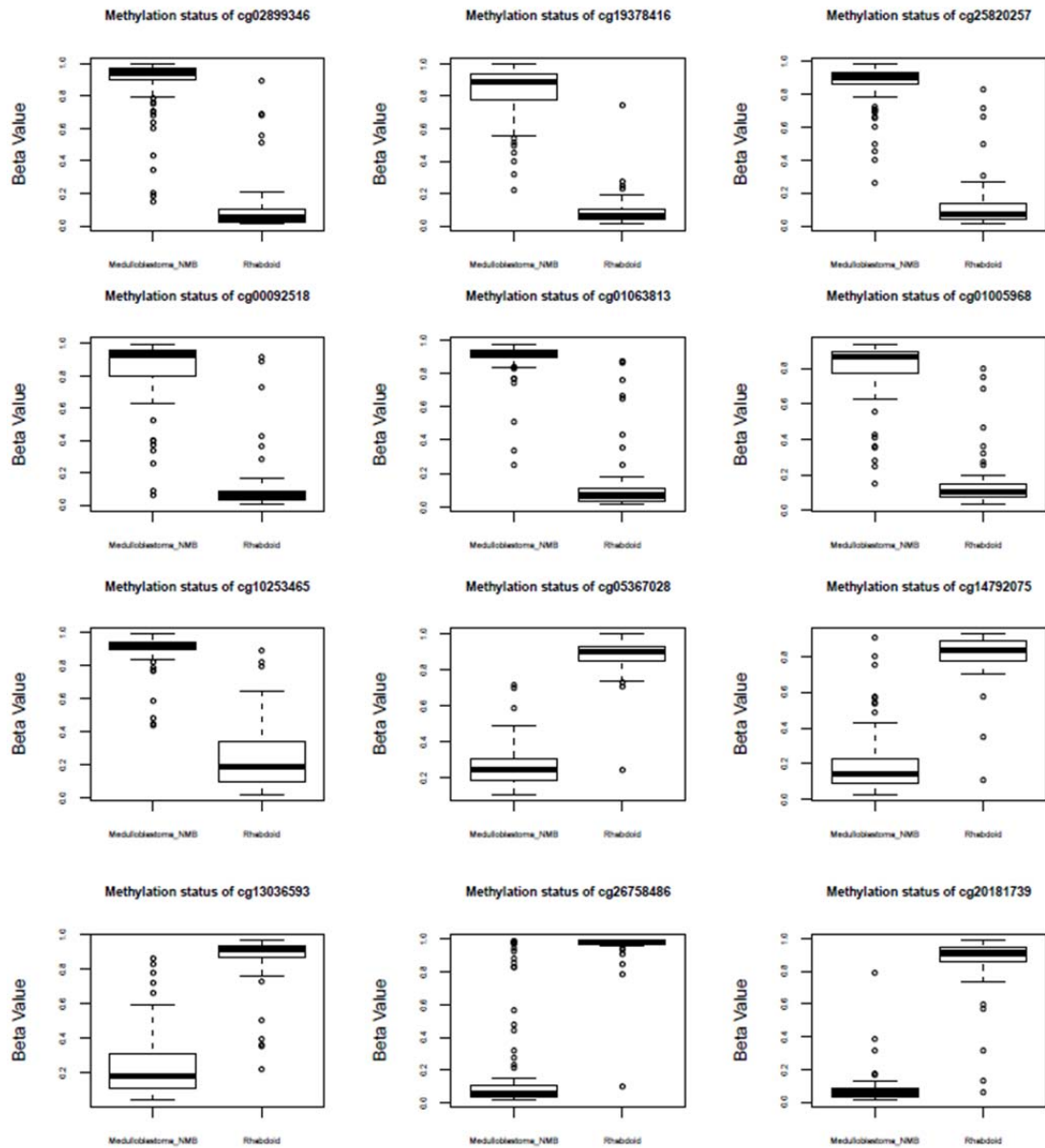


Figure 5.43 Relative difference in methylation status of 12 of 20 most deregulated genes between Rhabdoid Tumours (RT) and Medulloblastomas (MB). Box plots show  $\beta$ -values across the samples; median score (thick black line) and inter-quartile ranges (extent of box).

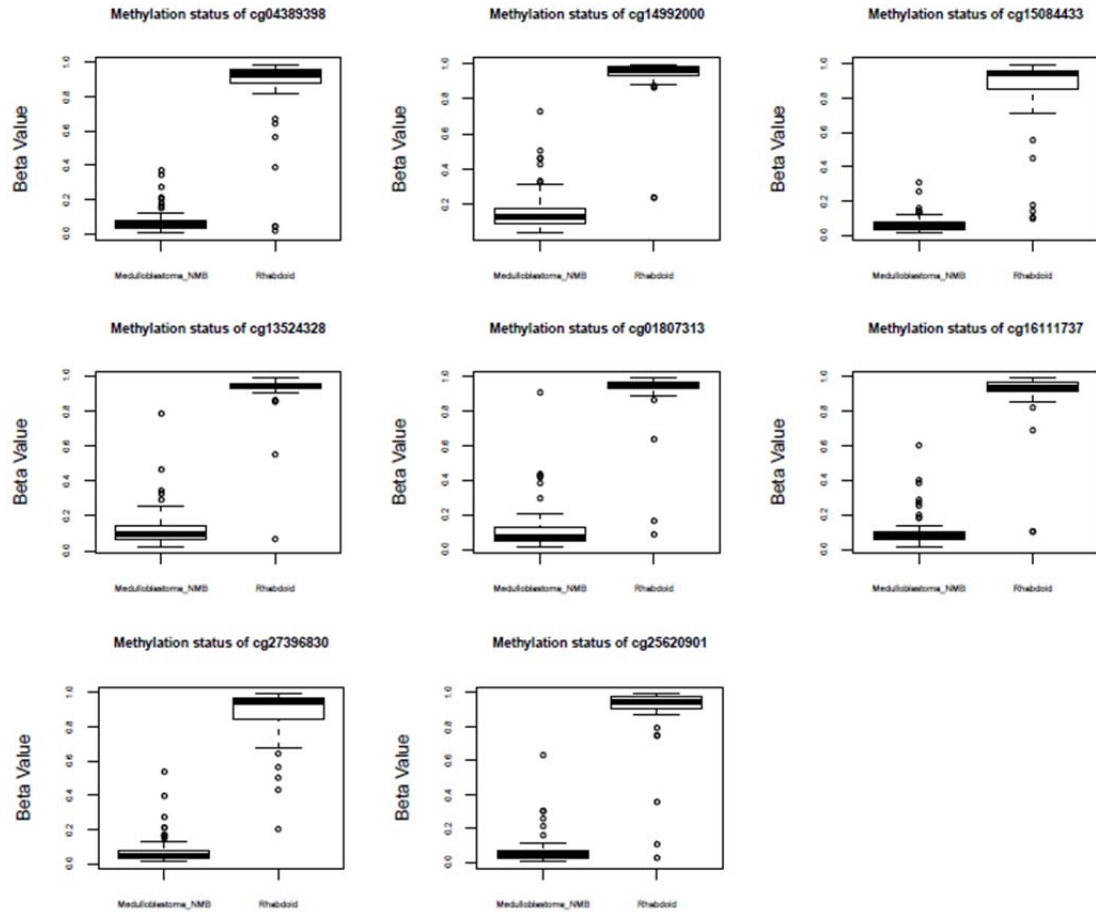


Figure 5.44 Relative difference in methylation status of 8 of 20 most deregulated genes between Rhabdoid Tumours (RT) and Medulloblastomas (MB). Boxplots show  $\beta$ -values across the samples; median score (thick black line) and inter-quartile ranges (extent of box).

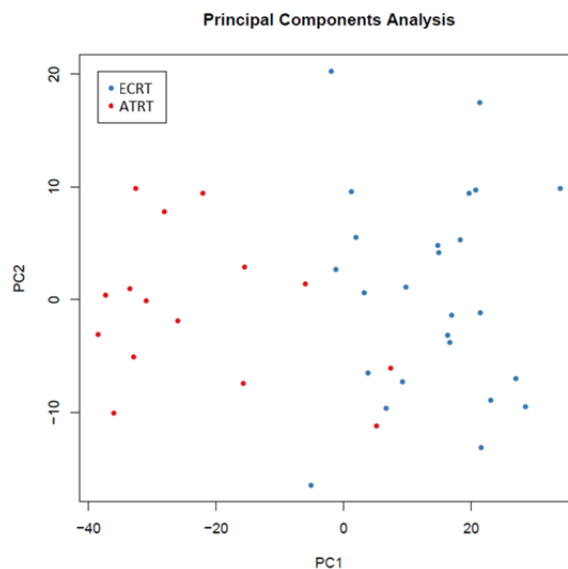
Many of the most significant alterations in methylation status occur in known oncogenes. The *PHF11* gene is associated with Ewing sarcoma and hypermethylation of this gene together with *NPTX2* is related with poor prognosis; also 36-50% of osteosarcomas showed methylation of *PHF11* gene (Alholle, Brini et al. 2013). In Neurofibromatosis type 1 down-regulation of *TNXB* affects cell–extracellular matrix interactions, eventually causing tumourigenesis (Levy, Ripoche et al. 2007); Variants of *TNXB* linked with *NOTCH4* variant gene were related with Age-related macular

degeneration. *FBXW8* controls cell cycle regulation promoting proteolysis of Cyclin D1; moreover *in vivo* studies showed its significant role in growth control: most *Fbxw8*<sup>-/-</sup> embryos die *in utero*, however a third showed abnormal development of the placenta (Okabe, Lee et al. 2006). Finally *DRAXIN* is involved in development of pre-cerebellar neurons, particularly it plays a repulsive role in axon outgrowth and nuclei migration (Riyadh, Shinmyo et al. 2014).

### 5.4.5 Rhabdoid Tumours present different DNA methylation subgroups mainly related to location of malignancy

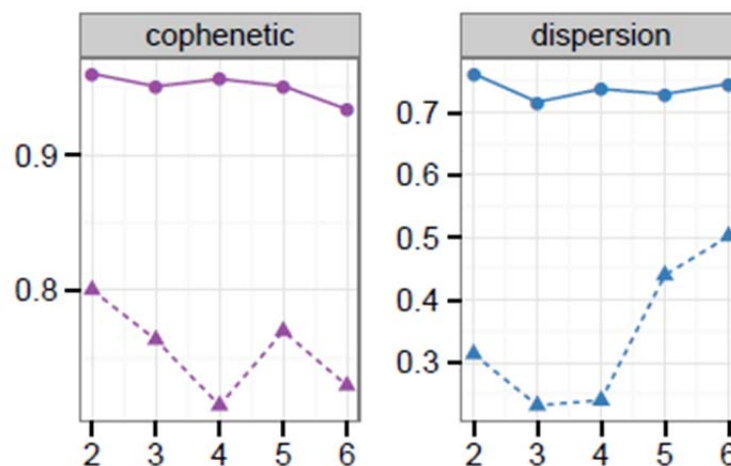
#### 5.4.5.1 Unsupervised analysis

Differences in methylation status between Extra Cranial Rhabdoid and Atypical Teratoid Rhabdoid Tumours were analysed using an unsupervised clustering approach (PCA, HC, and NMF). Principal component analysis (PCA) was performed using a subset of 936 probes ( $p < 0.01$ ) chosen as the CpGs which showed the greatest amount of variability across samples (Delta value  $< 0.5$  and  $> 0.5$ ). The PCA plot shows the tendency of that data to cluster into two biological groups, mainly although not exclusively related to the location of the tumours at the diagnosis (Figure 5.45).



*Figure 5.45 Unsupervised principal component analysis on 450K methylation data set characterises two subgroups within Rhabdoid Tumours. Points in red represent Atypical Teratoid Rhabdoid Tumours (ATRT) whilst Extra Cranial Rhabdoid Tumours (ECRT) are indicated in blue. The x-axis is the first component; the y-axis is the second principal component.*

A consensus cluster (repetitions =100) analysis was Initially carried out to identify the optimal number of metagenes and clusters. The stability of the cluster was determined by evaluating the cophenetic correlation and dispersion coefficient, as previously described. Metagenes were calculated for ranks  $k= 2$  to  $k=6$ ,  $k= 2$  was chosen as the most consistent subgrouping according to cophenetic correlation and dispersion coefficient (Figure 5.46). Consensus NMF and unsupervised hierarchical clustering identifies two subgroups mainly related to the location of the malignancies (Figure 5.47) with some exceptions for instance 2 Atypical Teratoid Rhabdoid Tumours grouped with the majority in Extra Cranial group and 4 in Extra Cranial Rhabdoid Tumours with the majority of the malignancies located in the brain .



*Figure 5.46 Cophenetic correlation and dispersion measures to assess the stability of clustering associated with each rank  $k$  for the methylation data set.*

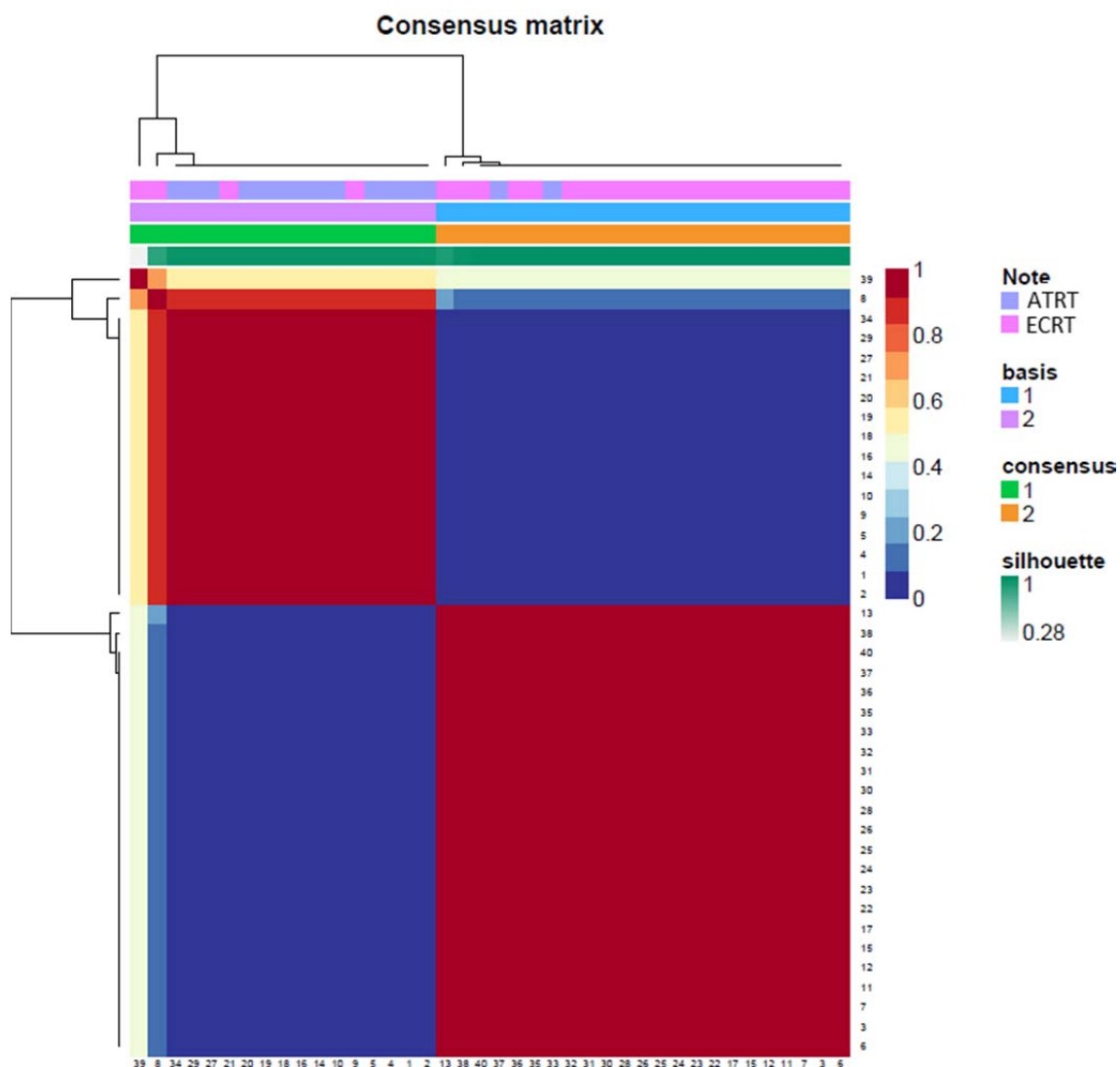
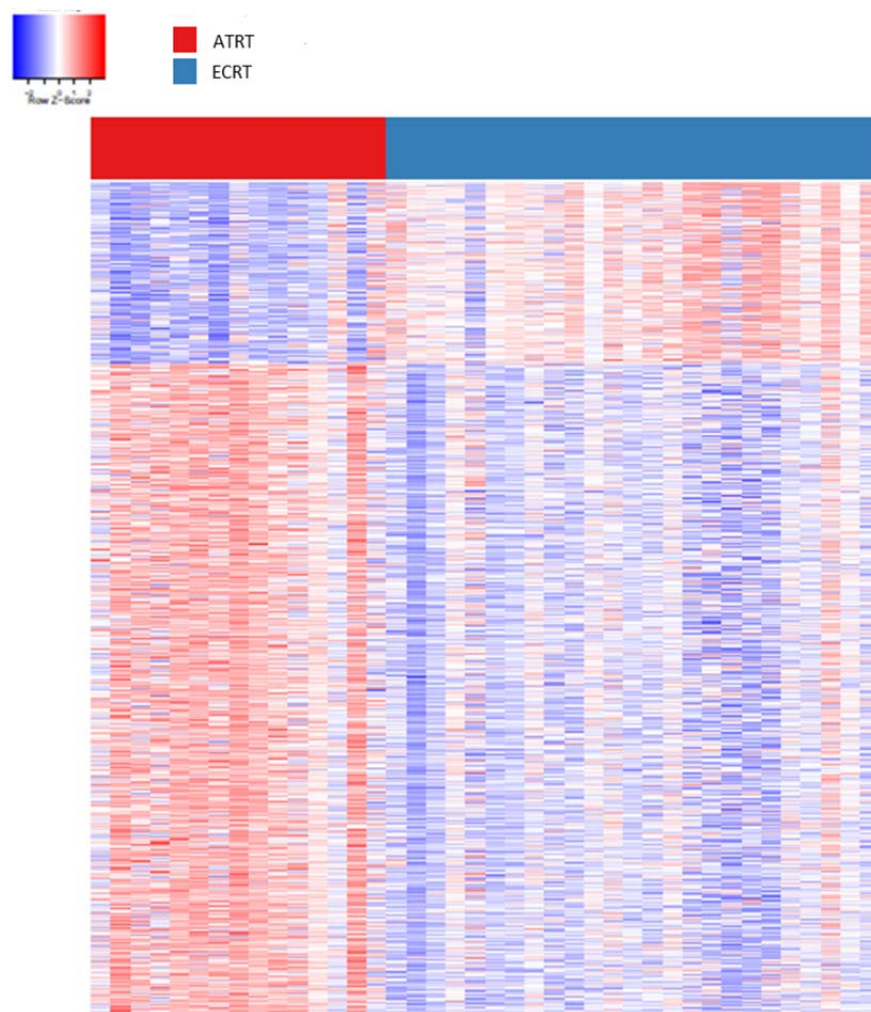


Figure 5.47 Consensus unsupervised NMF and hierarchical cluster analysis ( $k=2$ ) on Rhabdoid Tumour derived 450K methylation data set. Primary Rhabdoid Tumours are clustered into two different classes, mainly related to tumour locations. The matrixes were coloured 0 (deep blue, indicating samples never in the same cluster) to 1 (dark red, indicating samples always in the same cluster).

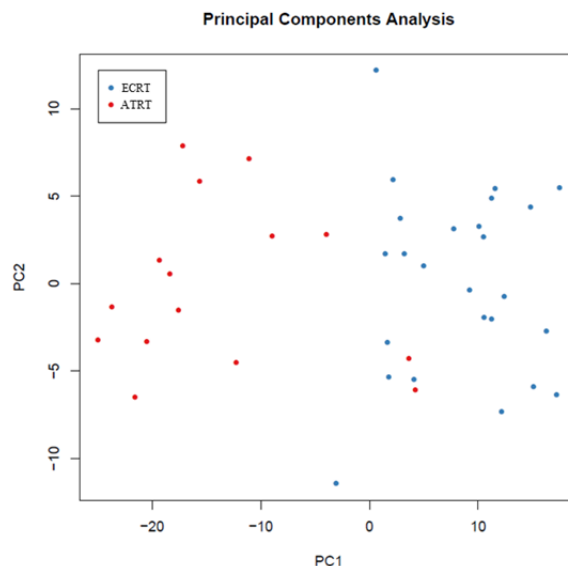
The heatmap Figure 5.48 in contains the top 100 most significantly differentially methylated probes (defined by p-value <0.01; delta value <0.5 and >0.5) and demonstrates the difference in methylation between Extra Cranial Rhabdoid and Atypical Teratoid Rhabdoid Tumours as well as those ATRT or ECRT which do not fall into their appropriate groups.



*Figure 5.48 Heatmap of gene expression showing the 100 most differentially methylated probes between Atypical Teratoid Rhabdoid Tumours (ATRT) and Extra Cranial Rhabdoid Tumours (ECRT). Each row represents a CpG probe whilst each column a sample. Hypermethylated (pale red to dark red) and Hypomethylated transcripts (pale blue to dark blue). In red are represent ATRT while ECRT are indicated in blue.*

#### 5.4.5.2 Supervised analysis

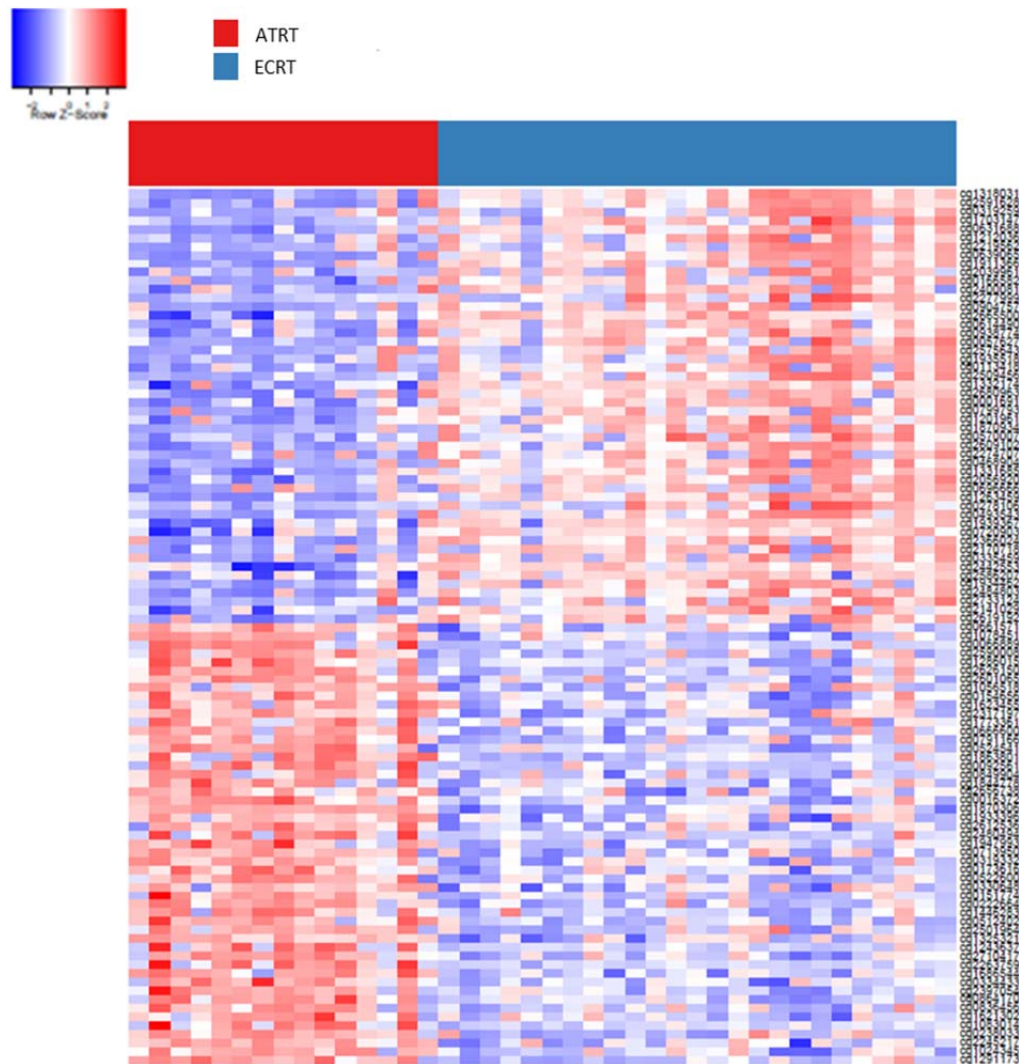
In order to judge the effect of selecting only those CpG probes which were significantly differentially methylated the 450K methylation data were filtered in a supervised fashion to create a subset composed only of CpGs which were significant differentially expressed. Specifically a PCA analysis (Figure 5.49) was performed using a subset of 316 probes ( $\beta$  value  $<-0.35$  and  $>0.35$ ;  $p<0.01$ ) and confirmed the tendency of that data to cluster in two biologically groups, mainly related to the location of the tumours at the diagnosis (Figure 5.42) but with two obvious exceptions where Atypical Teratoid Rhabdoid Tumours samples have grouped with in Extra Cranial Rhabdoid Tumours.



*Figure 5.49 Supervised principal component analysis on Malignant Rhabdoid Tumour 450K methylation data set distinguishes two biological groups, related to the tumour location. Points in red represent Atypical Teratoid Rhabdoid Tumours (ATRT) and Extra Cranial Rhabdoid Tumours (ECRT) are indicated in blue. The x-axis is the first component; the y-axis is the second principal component.*



The heatmap in Figure 5.50 contains the top 100 differentially methylated probes between the two different methylomic sub-groups found with the Rhabdoid Cohort and shows clear differences in methylation status between the two groups as well as the more obvious discrepancies between tumour location and methylation group.



*Figure 5.50 Heatmap of gene expression showing the most differentially methylated probes in Rhabdoid Tumours in supervised fashion. Each row represents a CpG probe whilst each column a sample. Hypermethylated (pale red to dark red) and hypomethylated transcripts (pale blue to dark blue). In red are represented Atypical Teratoid Rhabdoid Tumours (ATRT) and Extra Cranial Rhabdoid Tumours (ECRT) are indicated in blue.*

The methylation data set was subsequently filtered (adjusted  $p$ -value greater than 0.0001 and delta beta-value less than -0.7 or greater than 0.7) in order to identify significant methylated probes that define the two methylomic subgroups. Table 5.7, Figure 5.51 and Figure 5.52 show the 20 most differentially methylated probes resulting from this filtering (full results in Appendix 2). Four CpG probes are related to genes known to be associated with tumour progression (\*), two with cell cycle (+), one indicated as a potent tumour suppressor genes (~), one described as an oncogene (-) and one as a diagnostic marker and potential therapeutic target for lung cancer (<http://www.genecards.org/>). Complete list in Appendix I.

Probes	Gene Symbol	Average Delta values	LogFold change	Adjusted <i>p</i> value
cg16865446	PAX6* ~	0.50883	4.20386	0.000105
cg02380334	CCDC177	0.490687	4.366351	0.000105
cg05124021	PTPRN2	0.480161	3.979788	0.000164
cg02226645	MIR124-2	0.478548	3.9015	0.000206
cg05226008	DLC1~	0.461443	3.837403	0.000452
cg07139509	CCDC177	0.452786	3.804627	0.000616
cg08641706	LINC01102	0.448127	4.256806	0.00011
cg13480937	ASTN1	0.445405	3.401604	0.001051
cg27260617	AGTR1*	0.437639	3.192388	0.001466
cg01587682	PAX6	0.43528	3.482271	0.001271
cg12634591	HOXD3*	-0.40403	-3.40501	0.000818
cg25094569	WT1*-	-0.41285	-3.57374	0.000574
cg13638420	WT1	-0.41796	-3.13113	0.00204
cg22968401	PIPOX	-0.42332	-3.13263	0.001507
cg00016913	PIPOX	-0.42385	-3.5515	0.000452
cg09047573	NME5*	-0.42387	-3.74244	0.000619
cg09387749	HOXD3	-0.42725	-3.69163	0.000987
cg19344626	NWD1	-0.43282	-3.34403	0.001337
cg13321745	MSI1#	-0.43886	-3.56569	0.000574
cg19113668	ST5*	-0.44959	-3.88558	0.000248

*Table 5.7 Differentially methylated probes in Extra-cranial vs Atypical Teratoid Rhabdoid Tumours (adjust pvalue <0.01; Average delta values  $\geq +0.4$   $\leq -0.4$ ). Functional annotation of genes relevant to cancer is shown in the table with \* for gene associated with tumour progression, + with cell cycle, ~ for tumour suppressor genes, - for oncogene and # for gene described as diagnostic marker and potential therapeutic target.*

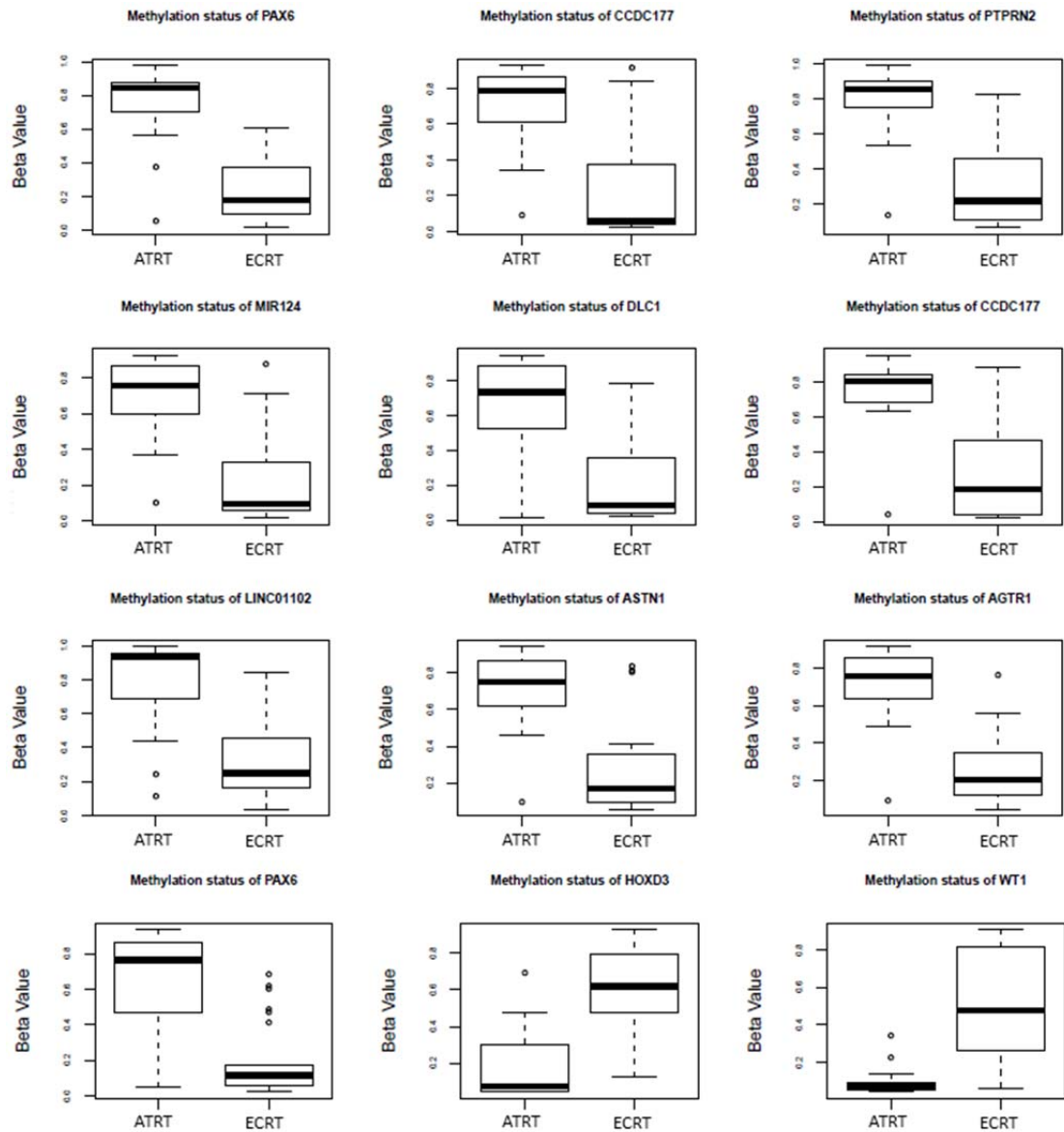


Figure 5.51 Relative difference in methylation status of 12 of the most deregulated genes between Atypical Teratoid Rhabdoid Tumours (ATRT) and Extra Cranial Rhabdoid Tumours (ECRT). Boxplots show  $\beta$ -values across the samples; median score (thick black line) and inter-quartile ranges (extent of box).

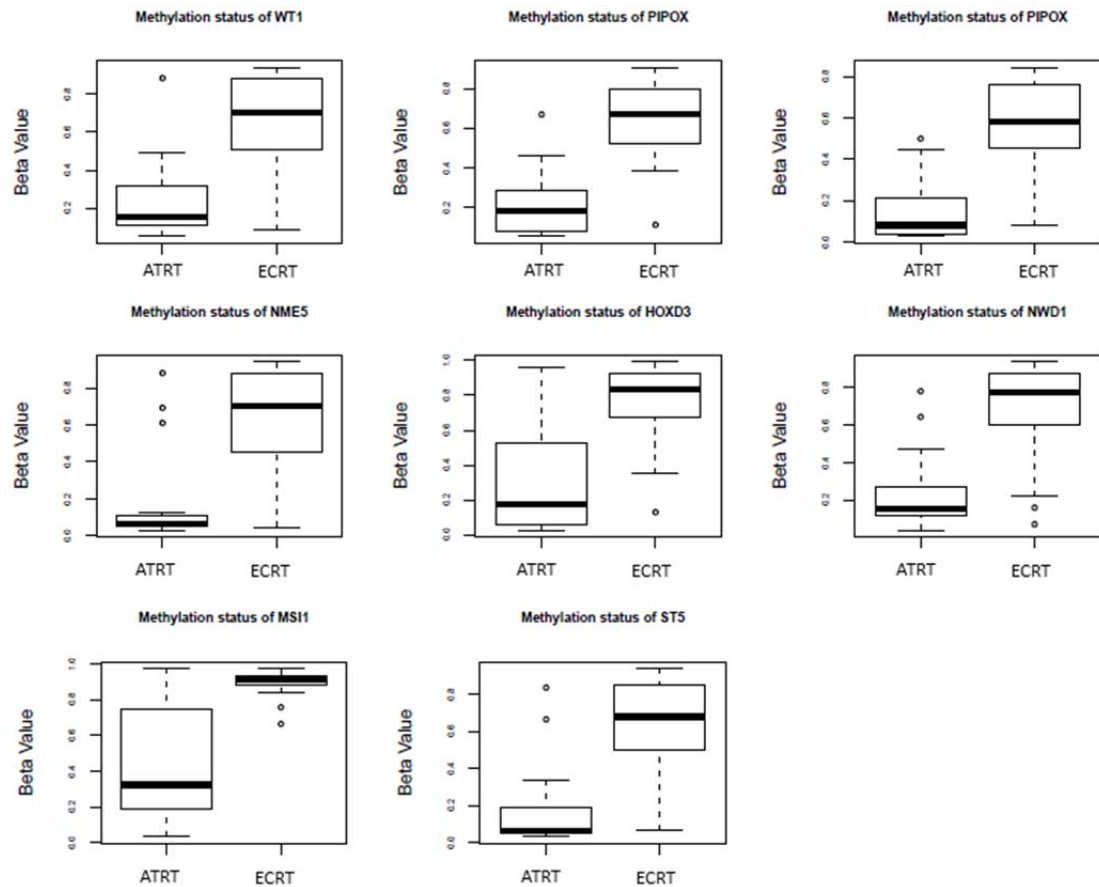


Figure 5.52 Relative difference in methylation status 8 of 20 most deregulated genes between Atypical Teratoid Rhabdoid Tumours (ATRT) and Extra Cranial Rhabdoid Tumours (ECRT). Boxplots show  $\beta$  values across the samples; median score (thick black line) and inter-quartile ranges (extent of box).

Many of the CpGs highlighted are associated with genes with a relevant role in tumorigenesis or development. For instance, *PAX6* is essential in the maintenance of neural stem and progenitor cells and plays oncosuppressive role in Oligodendroglioma (Appolloni, Calzolari et al. 2012). *DLC1* is a tumour suppressor gene with implication in a broad range of human cancer; it encodes for RhoGAP and therefore regulates kinases such as Akt, PKC and PKD (Ko, Chan et al. 2013). *HOXD3* is part of the homeobox genes, which play a key role in regulation of cardio-vasculature development and are expressed in normal placental development and essential in human foetal growth (Murthi 2014). *WT1* have been found fused with *EWS* in Desmoplastic small round-cell tumours; this fusion creates a tumourigenic chimeric protein which promotes aberrant neuronal differentiation (Kang, Park et al. 2014). *Msi1* is overexpressed in Lung cancer and suggested as a diagnostic marker (Wang, Yu et al. 2013); interestingly this gene directly regulates Wnt and Notch signalling pathways and is involved in neuronal development *in vivo* (Wang, Yin et al. 2008, Lan, Yu et al. 2010, Rezza, Skah et al. 2010).

## 5.5 Discussion

This chapter employed novel and extensive profiling technology for the characterisation of Rhabdoid Tumours. Rhabdoid Tumours were compared with paediatric Medulloblastoma malignancies to identify gene signatures that might explain pathogenesis. Because of the similar location of the mass at diagnosis Rhabdoid Tumours of the brain (ATRT) it is noteworthy that they are still misdiagnosed as Medulloblastoma tumours. In addition, analysis was carried out on the Rhabdoid Tumour Cohort alone to ask whether the biology driving Rhabdoid tumourigenesis differs according to different location in the body.

### 5.5.1 Malignant Rhabdoid Tumours are a unique biological group, with distinct methylation and expression pattern

Unsupervised gene expression analysis using RNA-seq of a Rhabdoid and Medulloblastoma cohort identified separate disease groups, underlining the clear difference in term of their molecular biology between these two tumours and that Rhabdoid Tumours as a whole form a single biological entity. Interestingly, consensus NMF analysis identifies a distinct gene pattern for Rhabdoid malignancies separate from all four Medulloblastoma subgroups (SHH, WNT and a heterogeneous group). Comparative analysis distinguished significant upregulation of *AMER2*, *HMGA2* genes and down-regulation of *ST18* and *NEUROG1* genes in Rhabdoid Tumours. *AMER2* is involved in Wnt signalling, a pathway which directs a wide range of developmental processes. *HMGA2* actively enhances *E2F*, which is transcriptionally activated by SMARCA4. *ST18* is a known tumour suppressor gene in Breast cancer and it has been suggested to be involved in Leukemogenesis, however the mechanism is still unknown. Methylation of *NEUROG1* causes tumourigenesis in uveal melanoma, concurrently with hypermethylation of *RASSF1A*.

Study of alternative usage of exons was performed to investigate the complexity of transcriptional processing in Rhabdoid Tumours. The comparative analysis against

Medulloblastoma identified 2002 genes with significantly differentially used exons, including *HES6*, *TUBA1C* and *RPS27L*. *HES6* is involved in Notch pathway, mechanism found deregulated in the paediatric soft tissue sarcomas; *TUBA1C* is generally overexpressed in Medulloblastoma cell line while *RPS27L* is involved in *P53* pathways.

Investigation of methylation status was carried out in Rhabdoid Malignancies using 450K methylation array. The analysis was performed by comparing Rhabdoid primary tumours with Medulloblastoma primary samples. Unsupervised analysis revealed two unique groups, one corresponding to Medulloblastomas and another to Rhabdoid Tumours, further confirming that Rhabdoid Tumours form a unique biological entity. Consensus NMF identified the same groups previously described in the RNAseq analysis. The investigation of specific methylomic patterns revealed amongst others hypermethylation of *PHF11*, *TNXB* and hypomethylation of *FBXW8* *DRAXIN* genes in Rhabdoid Tumours. *TNXB* is down-regulated in Neurofibromatosis and also linked to NOTCH pathway, however the tumorigenic mechanism is still unknown. *FBXW8* and *DRAXIN* play significant roles in cell growth, in embryonal and neuronal development respectively.

These findings are partially discordant with previously classification of Rhabdoid Tumours. The WHO organisation catalogues these malignancies as two different entities both characterised by *SMARCB1* deletion; this classification is still followed in clinics where Atypical Teratoid Rhabdoid Tumours and Extra Cranial tumours are treated with different regimen. Moreover, Birks et al. showed that Extra Cranial Rhabdoid and Atypical Teratoid Rhabdoid Tumours when compared with other CNS tumours shared a restrict number of deregulated genes. On the contrary we show that Rhabdoid Malignant Tumours are a unique and characteristic entity sharing many commonly expressed genes and methylated CpGs, with a very different biology from other CNS paediatric tumours.



### **5.5.2 Gene expression and methylation status reveals subgroups in Rhabdoid Malignant Tumours, mainly related to tumour location**

Further investigation of the Rhabdoid primary cohort were undertaken to identify biological subgroups. Unsupervised analysis of the RNA-seq gene expression data distinguished two separate subgroups, one mainly but not exclusively consisting of in Extra Cranial Rhabdoid Tumours and one mainly consisting of Atypical Teratoid Rhabdoid Tumours. Consensus NMF analysis confirmed two distinct gene patterns, separating the cohort into two metagenes related to the location of the primary tumours at diagnosis. Metagene projection of the results from our cohort onto three different data set acquired using Affymetrix array confirmed and validated the difference between Extra Cranial and Atypical Teratoid Rhabdoid Tumours as described by our cohort.

Gene expression analysis identified a significant subgroup of differentially expressed genes related to tumour progression, invasion and angiogenesis. Expression of genes involved in neurodevelopmental proliferation, such as *GFAP*, *B-FABP* and *SOX1*, were up-regulated in Atypical Teratoid Rhabdoid Tumours. Upregulation of *GFAB* is often detected in Glioblastoma and Chordomas as well as in Glioma (Theurillat, Hainfellner et al. 1999, Santos, Carvalho et al. 2009), where B-FABP family is also deregulated (Godbout, Bisgrove et al. 1998). SOX family proteins have a key role in embryonic and postnatal development and lead tumourigenesis in a wide number of cancers; in Hepatocellular Carcinoma hypermethylation of *SOX1* causes abnormal activation of Wnt/ $\beta$ -catenin signalling pathway and is a marker for cortical neural progenitor cells (Elkouris, Balaskas et al. 2011, Tsao, Yan et al. 2012).

Differential exon usage also indicates differences between the in Extra Cranial Rhabdoid Tumours and the Atypical Teratoid Rhabdoid Tumours. The analysis characterises 561 genes with significant differential exon usage in genes involved in Wnt signalling. *DCLK1* and *RNF220* genes are linked with regulation of the Notch and Wnt pathways; the relationship between these two genes is unknown. Wnt pathway is also deregulated by *HOPX* in lung tumour by and increased expression of *Mycn* (Pacheco-Pinedo, Durham et al. 2011).

Analysis of the methylation status of the Rhabdoid cohort confirms the presence of two methylomic defined subgroups mainly related to tumour location. Relative hypermethylation in Extra Cranial Teratoid Tumours was detected in CpGs associated with *PAX6* and *DLC1* (Appolloni, Calzolari et al. 2012); these genes play an oncosuppressive role in Oligodendroglioma. In Atypical Teratoid Rhabdoid Tumours *HOXD3* and *Msi1* are relatively hypermethylated: *HOXD3* is essentially expressed in normal placental development while *Msi1* directly regulates the Wnt and Notch signalling pathways (Murthi 2014).

Taken as a whole these analyses of Rhabdoid primary tumour expression and methylation profiles demonstrate the fact that even though these malignancies share a common biology when compared with other paediatric tumours, they differ significantly in their gene expression and methylation status. In particular, Atypical Teratoid Rhabdoid Tumours show a deregulation of a variety of genes involved in neurogenesis and neural-differentiation, whilst genes that define in Extra Cranial Rhabdoid Tumours have been differentially expressed in a variety of tumours without a clear link to a specific tissue. Interestingly, in Atypical Teratoid Rhabdoid Tumours a deregulation and change in methylation of genes regulating Wnt and Notch signalling was observed, indicating that this pathway might play an important role on the tumourigenesis of these malignancies

## **Chapter 6      Establishing the Role of *SMARCB1* loss in Malignant Rhabdoid Tumours**

## 6.1 Introduction

Malignant Rhabdoid Tumours present a very simple biology, since the only consistent mutation detected is loss of the *SMARCB1* gene and its respective encoded protein. Previous work, has demonstrated that *SMARCB1* re-expression induces a reversion of the tumorigenic phenotype of the cells, causing cell senescence by blocking cell cycle in G1. The effect of *SMARCB1* on gene expression has been previously studied by a number of people (Medjkane, Novikov et al. 2004, Gadd, Sredni et al. 2010); however the understanding of this mechanism was limited by the number of the cell lines and by the techniques used. In these studies microarray experiments were carried out in MON, STA-WT1, G401 cell lines upon *SMARCB1* re-expression and in a *SMARCB1* murine cancer model cells (Guidi, Sands et al. 2001, Medjkane, Novikov et al. 2004, Morozov, Lee et al. 2007, Gadd, Sredni et al. 2010). Just one study was focused on exploring the downstream effect of *SMARCB1* re-expression, while other studies were more focused on a better understanding of previously described pathways. In particular, microarray analysis highlights the effect of *SMARCB1* re-expression on gene expression and in particular how it causes up-regulation of several cell cycle control genes such as *CDK1A* and *CDK2A*, involvement of cell growth and actin cytoskeleton organisation genes such as *E2F* and *Rho-GAP4* (Medjkane, Novikov et al. 2004), re-expression of senescence genes such as *SERPINE2* (Gadd, Sredni et al. 2010) and down-regulates mitotic genes such as *TOP2A*, and *Aurora A* (Morozov, Lee et al. 2007). However, these investigations did not attempt to address the genome-wide effects of *SMARCB1* loss in order to understand downstream effects as well as to identify therapeutically targetable genes.

This part of the study was focused on studying the genomic effect of *SMARCB1* re-expression in Rhabdoid cell models. Gene expression analysis using RNA-seq together with DNA methylation analysis by 450K methylation array was performed for the first time in four *SMARCB1* stable infected Rhabdoid cell lines. Moreover, to understand *SMARCB1* function in methylation, cell lines treated with the demethylating agent 5-aza-2'-deoxycytidine were also subjected to sequencing by RNA-seq. The bioinformatics analysis and data integration pursued in this study allowed me to

uncover the general scale and nature of genomic and epigenetic aberrations caused by *SMARCB1* loss in Rhabdoid Tumours.

## **6.2 Aims**

The study reported in this chapter aimed to:

- Assess the effect of *SMARCB1* re-expression on gene expression and methylation status in Rhabdoid cells derived from different location.
- Explore gene expression changes in Rhabdoid cells following treatment with 5-aza-2'-deoxycytidine.

### **6.3 Integrated genomic and epigenetic analysis identifies downstream target genes dependent on *SMARCB1***

For the first time a genome-wide analysis combining expression analysis with methylation profiling was carried out in Rhabdoid cell lines to collect an extensive catalogue of late effects of *SMARCB1* re-expression on gene regulation and methylation status. In this analysis the previously generated (see Chapter 4) *SMARCB1* positive cell lines were analysed using an RNA-seq and 450K methylation approach.

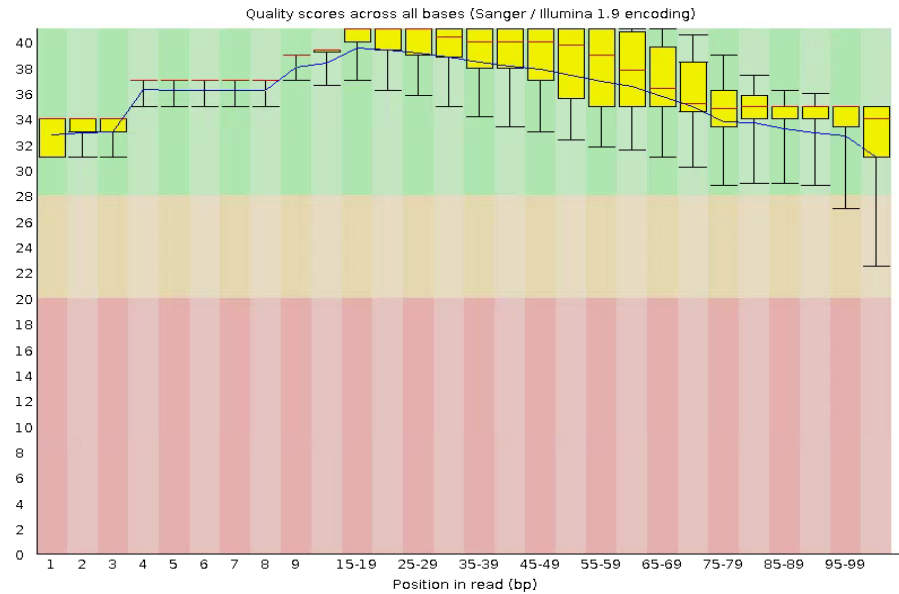
#### **6.3.1 RNA-seq analysis**

##### *6.3.1.1 Primary samples preparation*

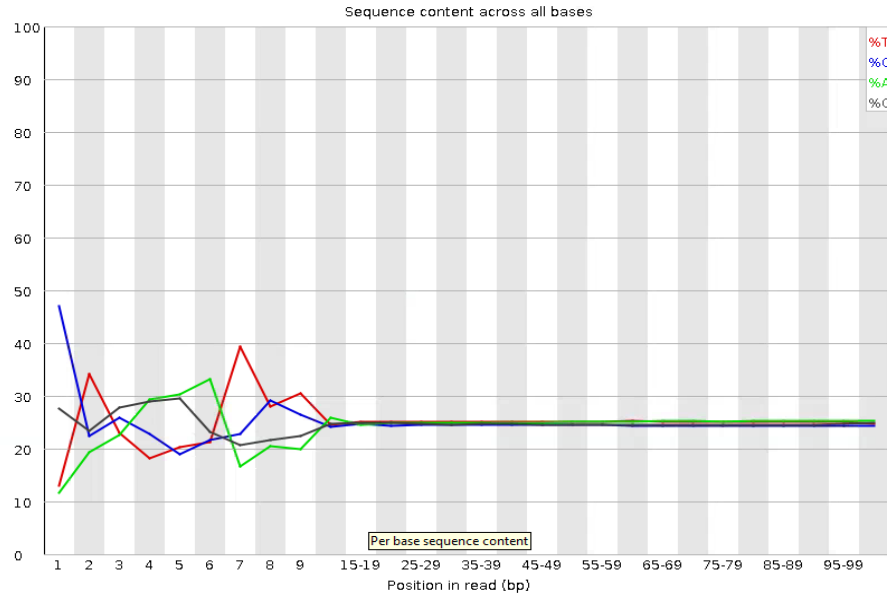
RNA from the cell was extracted as indicated in section 1.1.1 RNA integrity number (RIN score) was evaluated for each sample, using a Bioanalyzer. All the RNA samples were scored with a RIN greater than 9.

##### *6.3.1.2 RNA-seq QC analysis*

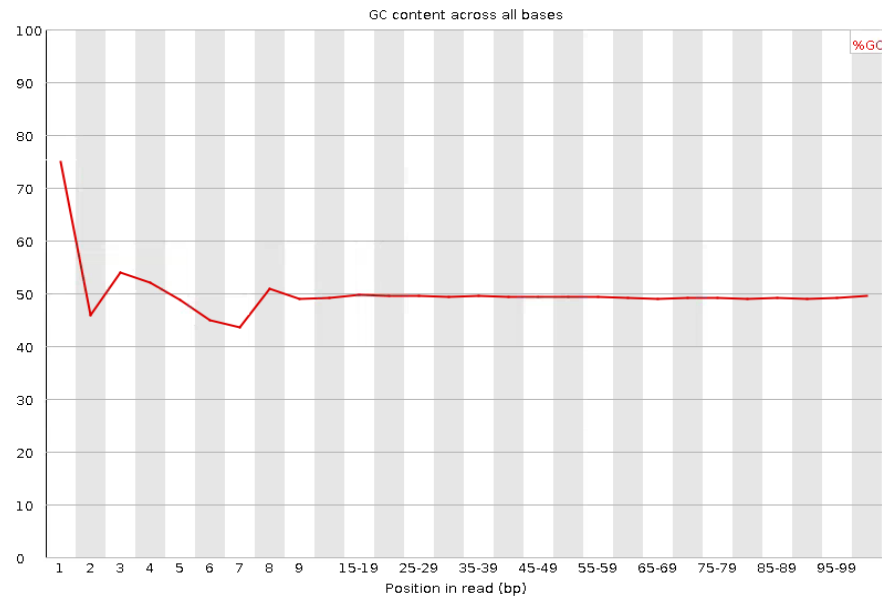
All samples sequenced successfully passed the RNA quality control checks, in particular in every sample the average Phred score across all bases was greater than 30 (Figure 5.2). As was the case in the sequenced primary sample, FASTQC analysis presented warnings for the per base sequence content, per base GC content (Figure 6.2 and Figure 5.4) and sequence duplication levels (Figure 6.4). This confirmed the high variability on the first bases as result of random priming.



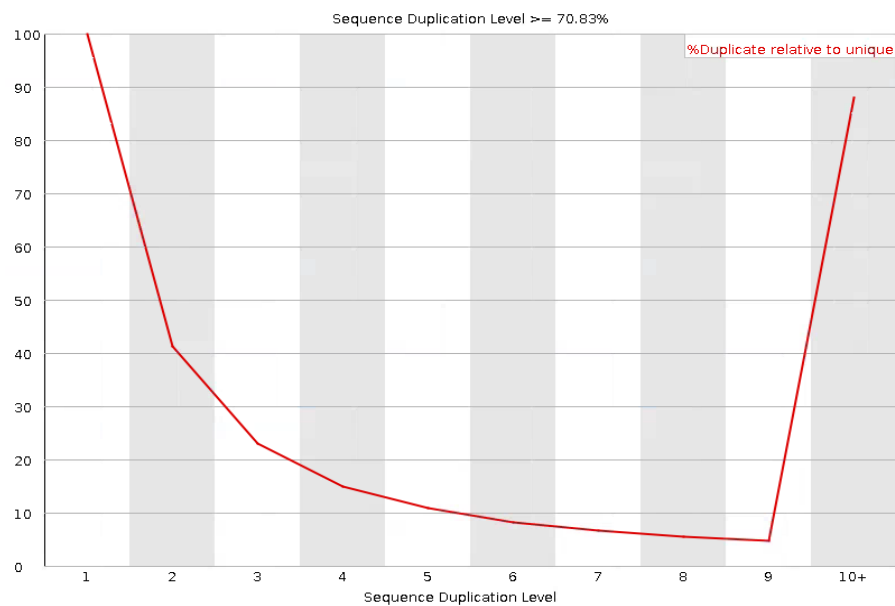
*Figure 6.1 Per base sequence quality derived from sample G401 SMARCB1+ from FASTQC html report. The red line represents the median value, while the blue line the quality. The plot presents 3 different background shading: green indicates a good quality score.*



*Figure 6.2 Per base GC content (B) of sample G401 SMARCB1+ from FASTQC html report. High variability of the first 10 bases was observed in all samples of our data set.*



*Figure 6.3 Per base sequence content analysis (A) and Per base GC content (B) of sample G401 SMARCB1+ from FASTQC html report. High variability of the first 10 bases was observed in all samples of our data set; this phenomenon is more plausibly constitutive biases rather than batch effects.*

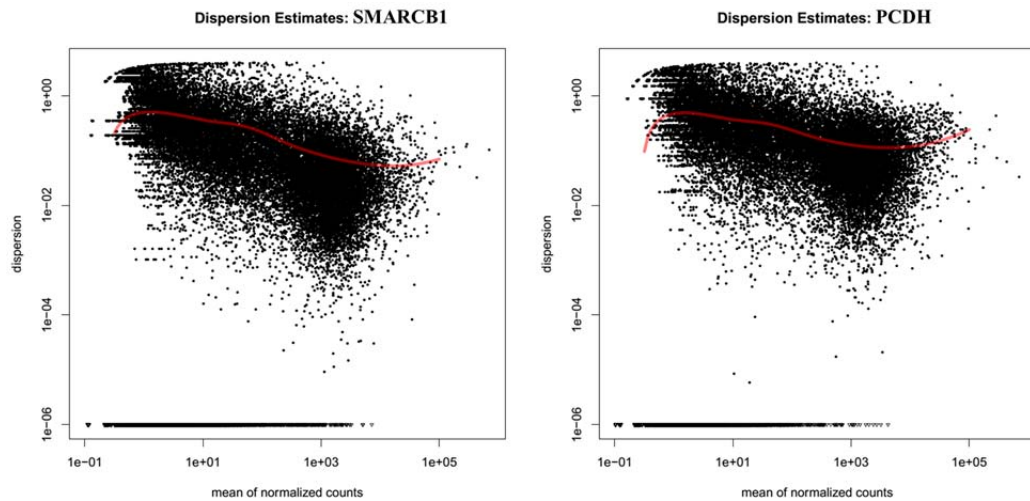


*Figure 6.4 Sequence duplication levels plot of sample G401 SMARCB1+ from FASTQC html report. The 10th column includes all duplication levels greater than or equal to 10.*

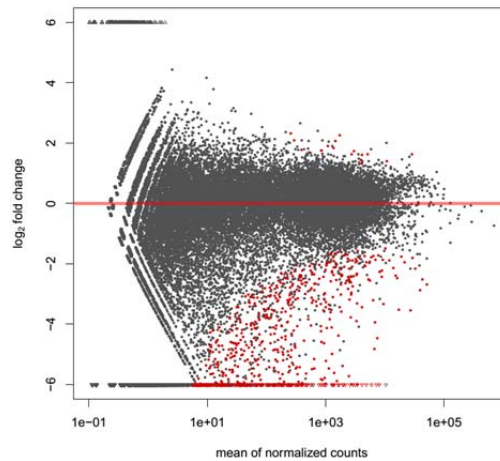


### 6.3.1.3 DESeq analysis

Raw count values were analysed in a pair wise fashion using DESeq2 in R. The negative binomial method was used to identify differentially expressed genes, as described in section 3.16. Estimates of dispersion in the data set are indicated in Figure 5.10, where the black dots represent the empirical dispersion values and the red line the fitted values. Figure 5.11 represents the differential expression estimation between *SMARCB1* re-expressing cell lines and the control infectants, where the red dots represents genes marked as differentially expressed at 10% FDR.



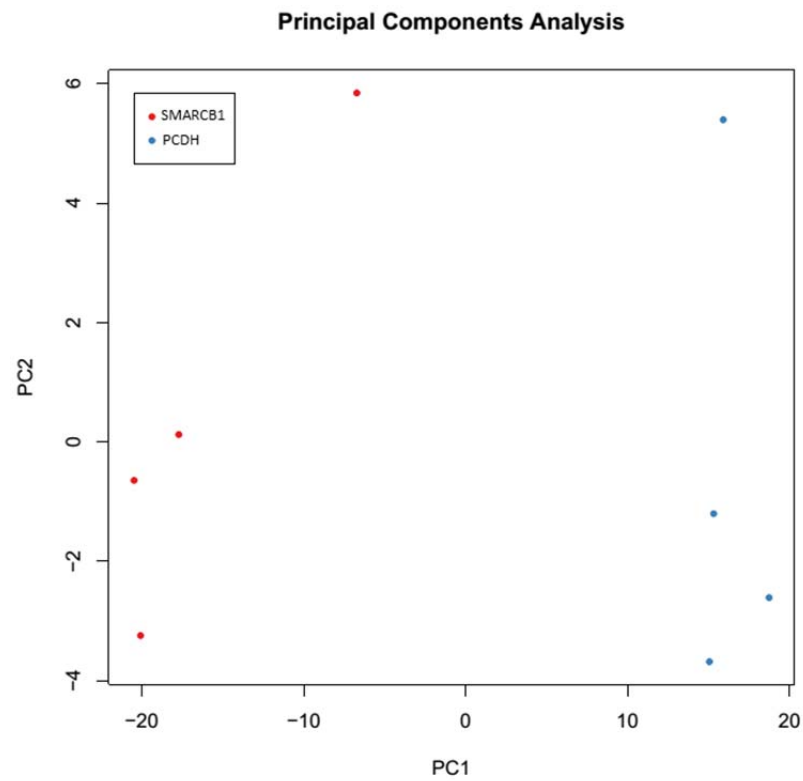
*Figure 6.5 Gene expression dispersion estimation and model fitting of SMARCB1 re-expressing cells (SMARCB1) and control infectants (PCDH). Each dot represents per gene dispersion estimation and the red line the fitted mean dispersion function.*



*Figure 6.6 Testing for differential expression between SMARCB1 re-expressing cells and control infectants. Log2fold versus mean expression over all samples. Red dots represent genes marked as differentially expressed at 10% FDR, when Benjamini and Hochberg correction is applied. The lines on the top and on the bottom of the plot represent genes with a very low or very high log fold change (value off scale).*

### *Clustering analysis*

DEseq results were filtered to subset only significant genes (adjusted  $p$ -value  $< 0.01$ ; moderated log fold change  $\leq -2$  and  $\geq +2$ ). PCA analysis was performed on this subset of 493 genes; the analysis identified two separate subgroups, one constituted by control infectants and the other cell lines re-expressing *SMARCB1* (Figure 6.7). Thus confirming the changing expression pattern and reflecting the change in phenotype of Rhabdoid cells following *SMARCB1* re-expression.



*Figure 6.7 Supervised principal component analysis on Rhabdoid cell lines expression distinguishes two different biological groups. Red points represent SMARCB1 re-expressing Rhabdoid cells (SMARCB1) and blue points vector control infected Rhabdoid cells (PCDH). The x-axis is the first component; the y-axis is the second principal component.*

Consensus NMF supervised analysis ( $k=2$ ) was carried out on the selected genes identifying two different metagenes; one for the *SMARCB1* positive cells and one for the control cells (Figure 6.8).

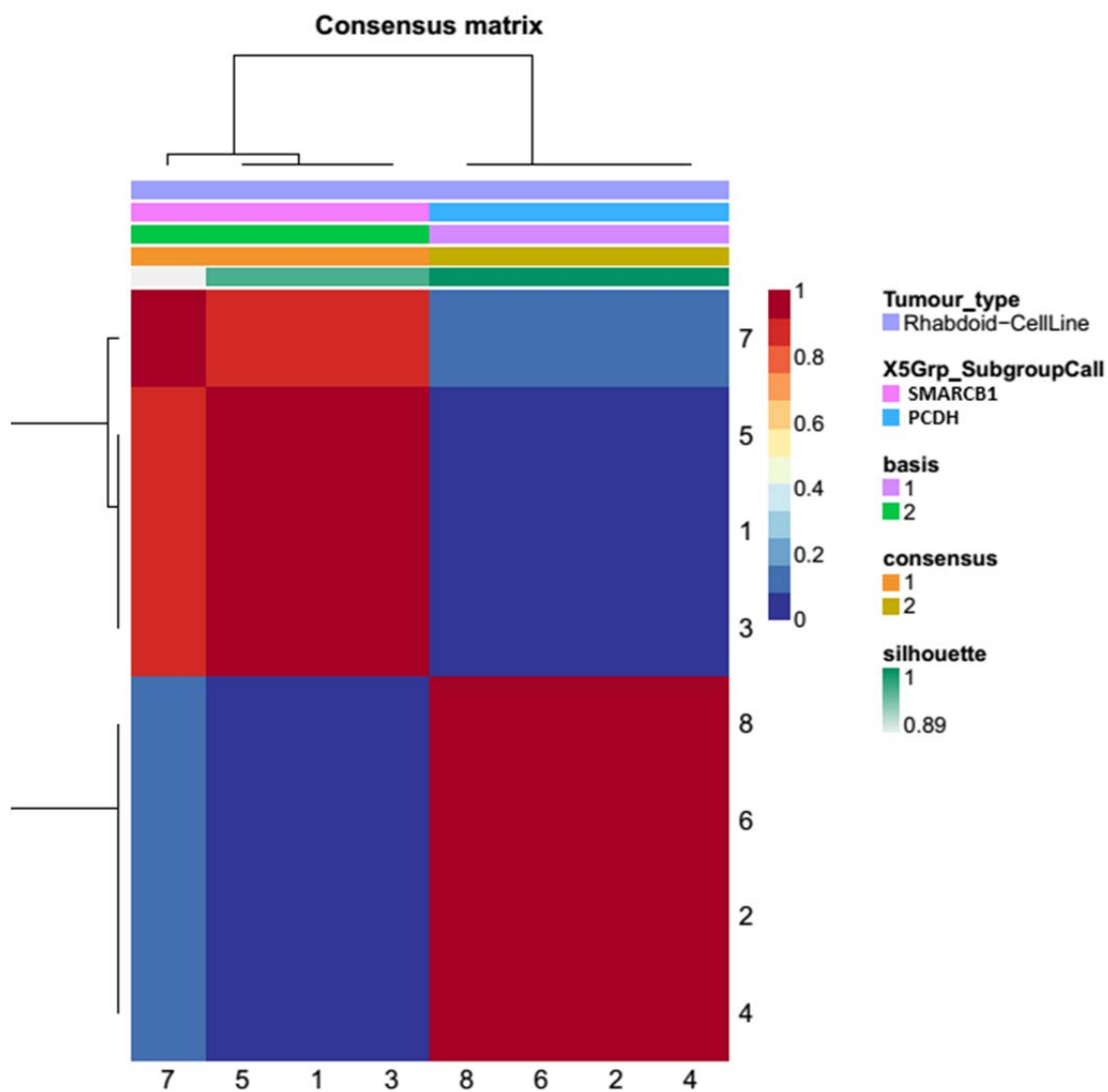
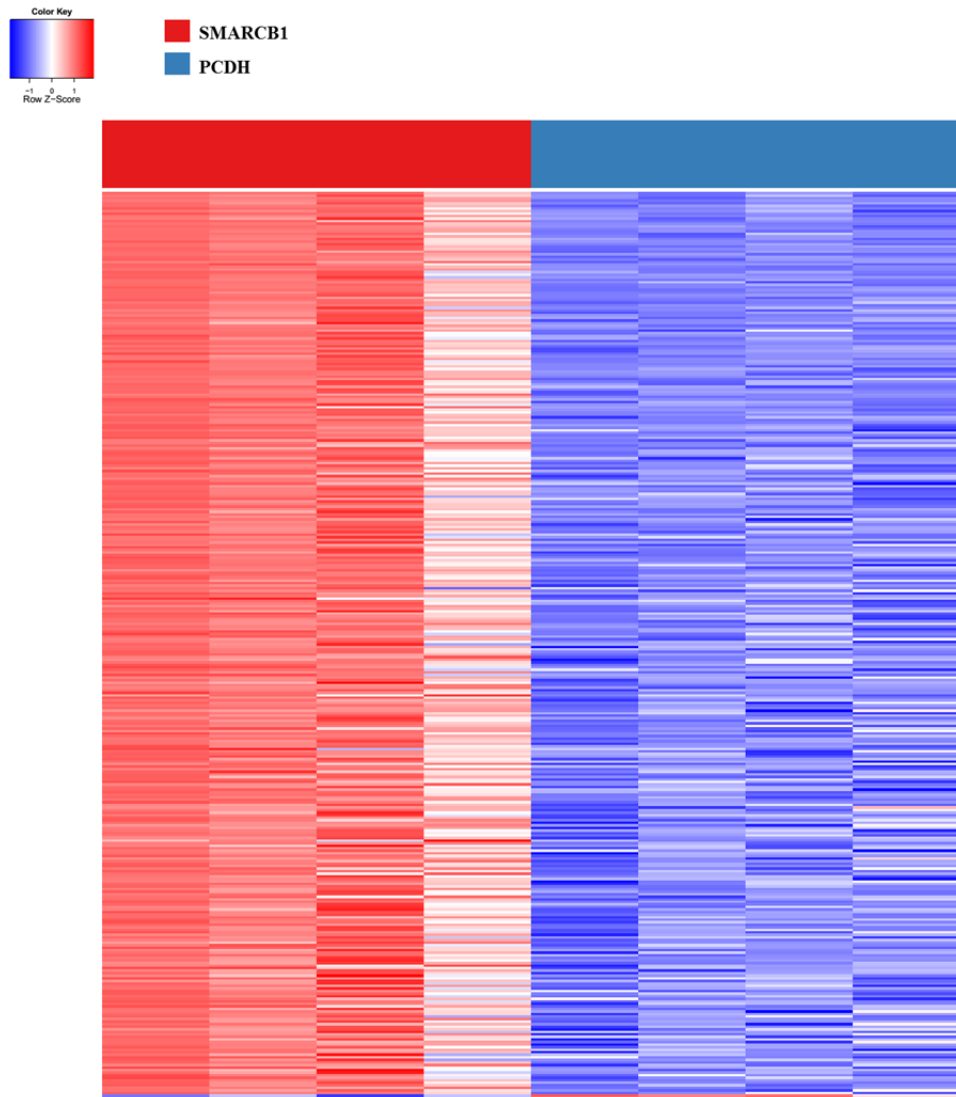


Figure 6.8 Supervised NMF and hierarchical cluster analysis ( $K=2$ ) on expression data from SMARCB1 re-expressing cells (SMARCB1) and control infectants (PCDH). The matrices were coloured 0 (deep blue, indicating samples never in the same cluster) to 1 (dark red, indicating samples always in the same cluster).

Filtered data were used to generate a heatmap with the top 100 deregulated genes (Figure 6.9) revealing that *SMARCB1* re-expression mainly causes significant up regulation of the majority of the genes analysed.



*Figure 6.9 Heatmap of significantly differentially expressed genes in SMARCB1 re-expressing Rhabdoid cells (SMARCB1) and in vector control infected Rhabdoid cells (PCDH). Each row represents a probes and each column a sample. Overexpressed (pale red to dark red) and under-expressed transcripts (pale blue to dark blue).*

In Table 6.1 are indicated the 15 most differentially expressed genes as a result of filtering, the relative difference in expression of these genes are indicated in Figure 6.10 and Figure 6.11. Seven genes are associated with tumourigenesis (\*), one instead is indicated as an antiangiogenic factor (#) two involved in cell proliferation (+), two directly involved in tumour invasion (-) (<http://www.genecrads.org>). Complete list in Appendix I.

Gene Symbol	Adjusted <i>p</i> -value	Moderated log2fold
<b>DNA2 *</b>	1.21E-19	7.552249108
<b>RBM20</b>	2.54E-49	6.976379538
<b>PLS1*</b>	2.62E-13	6.69712796
<b>DBP</b>	4.33E-18	6.333667579
<b>MSX2*-</b>	4.13E-18	6.302317536
<b>MMRN1</b>	3.90E-34	6.249127503
<b>LBX1-+</b>	5.21E-38	5.883944727
<b>ABHD2</b>	1.16E-29	5.725973275
<b>EMP2 *-</b>	3.30E-37	5.426349508
<b>CDC25C*</b>	1.91E-07	5.187714389
<b>GNG4</b>	3.67E-22	5.184757846
<b>SERPINF1 #</b>	1.90E-21	5.161774575
<b>CDC42BPG*</b>	0.000397	-1.602650115
<b>ENTPD7</b>	6.01E-08	-1.848080725
<b>SLC16A9* +</b>	2.44E-06	-2.071080896

*Table 6.1 Selected differentially expressed genes in SMARCB1 re-expressing Rhabdoid cells vs vector control infected Rhabdoid cells. Functional annotation of genes relevant to cancer is shown in the table with \* for genes associated with tumourigenesis, + for genes involved in cell proliferation, - for genes directly correlated with tumour invasion, # for genes indicated as inhibitors of tumourigenesis and tumour invasion.*

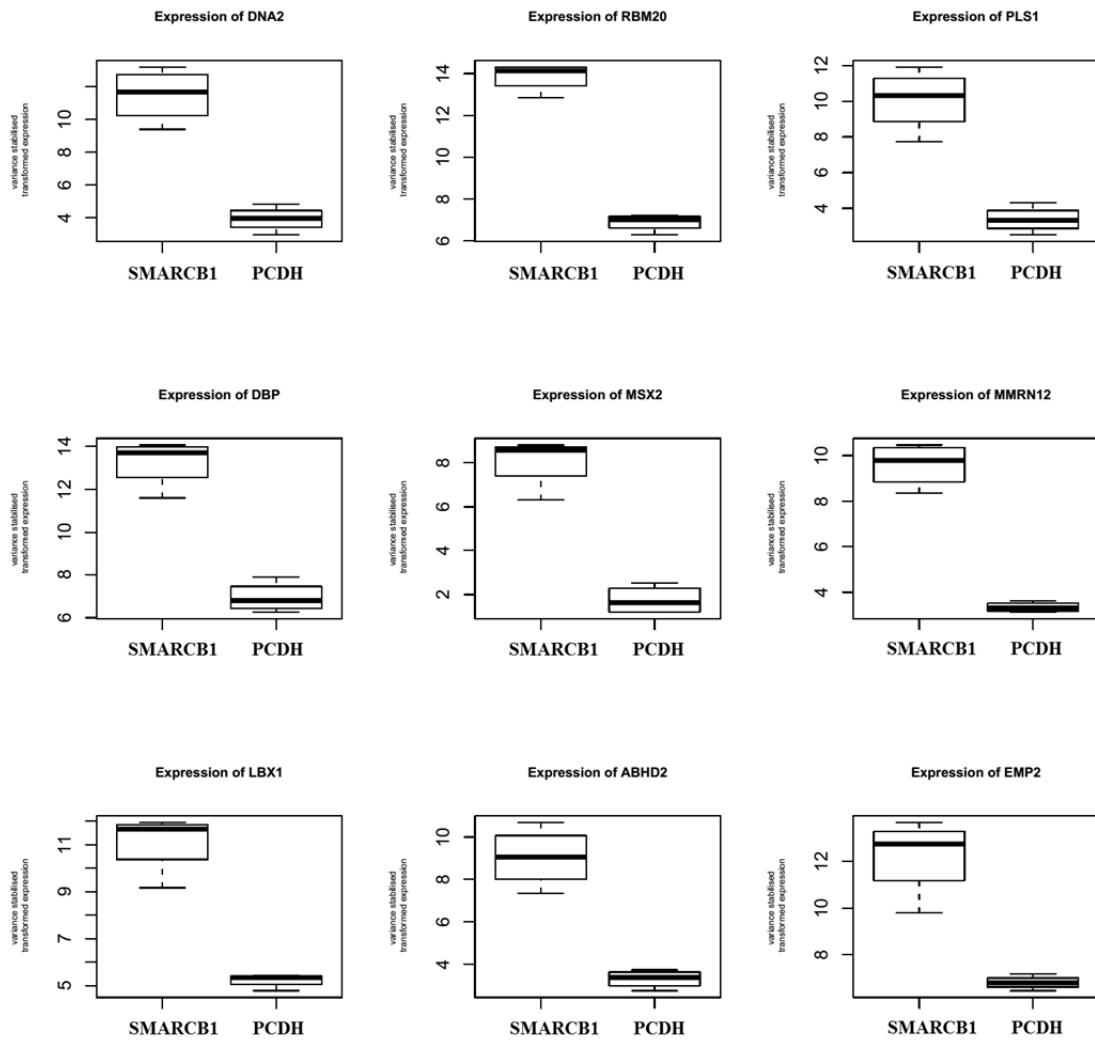


Figure 6.10 Relative difference in expression of 9 of 15 most deregulated genes between SMARCB1 re-expressing Rhabdoid cells (SMARCB1) and vector control infected Rhabdoid cells (PCDH). Boxplots show variant stabilised transformed expression across the samples. Boxplots show median score (thick black line) and inter-quartile ranges (extent of box).

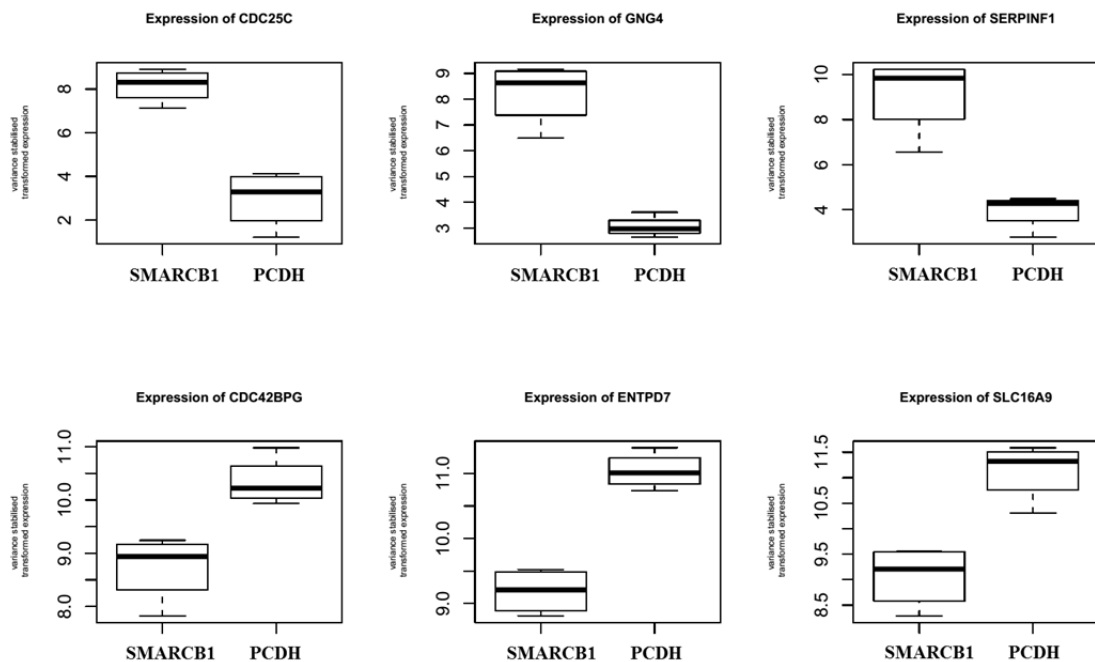


Figure 6.11 Relative difference in expression of 6 of 15 most deregulated genes between SMARCB1 re-expressing Rhabdoid cells (SMARCB1) and vector control infected Rhabdoid cells (PCDH). Boxplots show variant stabilised transformed expression across the samples. Boxplots show median score (thick black line) and inter-quartile ranges (extent of box).

DNA2 encodes for the human nuclease/helicase DNA2 protein, involved in repair replication-associated DNA double-strand breaks (DSBs). Depletion of the DNA2 causes chromosomal aberrations and inter-nuclear chromatin bridges, affecting nuclear DNA stability (Duxin, Dao et al. 2009). *SERPINF1* is involved in a variety of biological process, including neuronal development (Tombran-Tink, Chader et al. 1991); deregulation of *SERPINF1* following SMARCB1 re-expression has also been previously described (Gadd, Sredni et al. 2010). *LBX1* is a developmentally regulated homeobox gene involved in neurogenesis; interestingly although *LBX1* is physiologically expressed during neuronal differentiation, a high expression level is associated with higher tumour grades in breast cancer (Yu, Smolen et al. 2009).



#### 6.3.1.4 DEXSeq analysis: differential exon usage

Raw count values were analysed to study alternative splicing and other forms of alternative isoform expression, using a negative binomial method i.e. DEXSeq (see method Chapter 3). In Figure 6.12 estimates of dispersion in the data are indicated, whereby the black dots represent the empirical dispersion values and the red line the fitted values; the MA plot shows the distribution of exon usage for each gene, comparing vector control cells and *SMARCB1* re-expressing Rhabdoid cells.

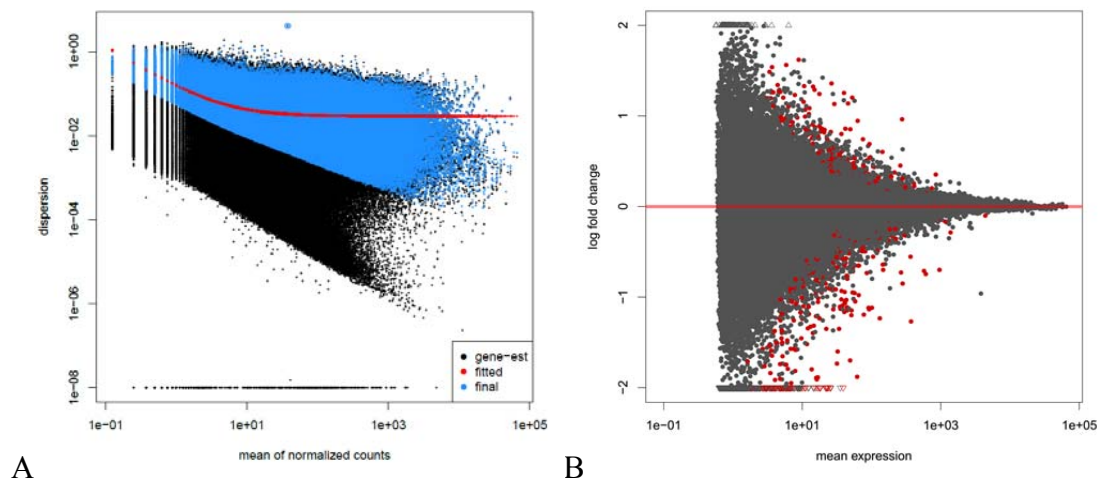


Figure 6.12 Per exon usage dispersion estimation and model fitting (A) and testing for differential exon usage of genes (B) comparing *SMARCB1* re-expressing cells and control infectants. A) Each dot represents per exon dispersion estimation and the red line the fitted mean dispersion function. B) Log2fold change versus mean expression over all samples. A red dot represents exons marked as differentially used at 10% FDR, when Benjamini and Hochberg method is applied. The lines on the top and on the bottom of the plot represent exons with a very low or very high log fold change (value off scale).

### 6.3.1.5 Detection of differentially used exon in gene *SMARCB1* dependent

The DEXseq results were filtered for values with an adjusted *p*-value less than 0.05 and log2 fold change less than -1 or greater than 1. Moreover, only protein coding genes were analysed; 59 genes with significantly differential exon usage were identified. Table 6.2 shows 15 genes with the most significantly differentially used exons following filtering. Of the 15 genes in the table three were associated with tumour progression (\*), one involved in autophagy (+) one described as tumour suppressor genes (~) and as a biomarker of poor outcome in Neuroblastomas (#) (<http://www.genecrads.org>). Complete list in Appendix I.

Gene Symbol	Exon	Adjusted <i>p</i> value	log2fold
<b>EMC10</b>	E009	1.23E-07	1.619989
<b>FAM60A*</b>	E015	1.25E-07	1.56494
<b>LAMB1</b>	E012	5.29E-08	1.53708
<b>ALG13#</b>	E036	2.37E-07	1.487411
<b>ERC2*</b>	E002	4.12E-09	1.331395
<b>SYNPO</b>	E012	5.90E-07	1.321838
<b>AP2S1</b>	E002	1.68E-11	1.28767
<b>NUMA1*</b>	E035	4.50E-08	-6.27551
<b>RUSC1</b>	E025	5.61E-07	-6.57386
<b>SCOC</b>	E024	2.95E-08	-6.71161
<b>DLD</b>	E027	8.48E-11	-7.01253
<b>MAPKAP1</b>	E013	3.11E-07	-7.58521
<b>PRKD2 *</b>	E009	5.48E-10	-10.8838
<b>BTG4 ~</b>	E020	3.26E-08	-11.2854
<b>PGAM1*</b>	E016	2.86E-12	-11.7591

*Table 6.2 Genes with differentially used exons in SMARCB1 re-expressing Rhabdoid cells vs control infected cells. Functional annotation of genes relevant to cancer is shown in the table with \* for genes associated with tumour progression, + indicates genes related to autophagy, ~ tumour suppressor genes (~) and as biomarker of worst outcome (#).*

*BTG4* is part of the anti-proliferative BTG family; down-regulation and hypermethylation of *BTG4* was observed in gastric cancer and its re-expression in cell lines causes tumour growth inhibition (Dong, Tu et al. 2009), however it is highly expressed in preimplantation embryos (Buanne, Corrente et al. 2000). Figure 6.13 shows the differential exon usage of *BTG4* obtained from DEXSeq analysis. Exonic part 1 is affected by *SMARCB1* re-expression in Rhabdoid cells indicating a probable alternative transcriptional start site.

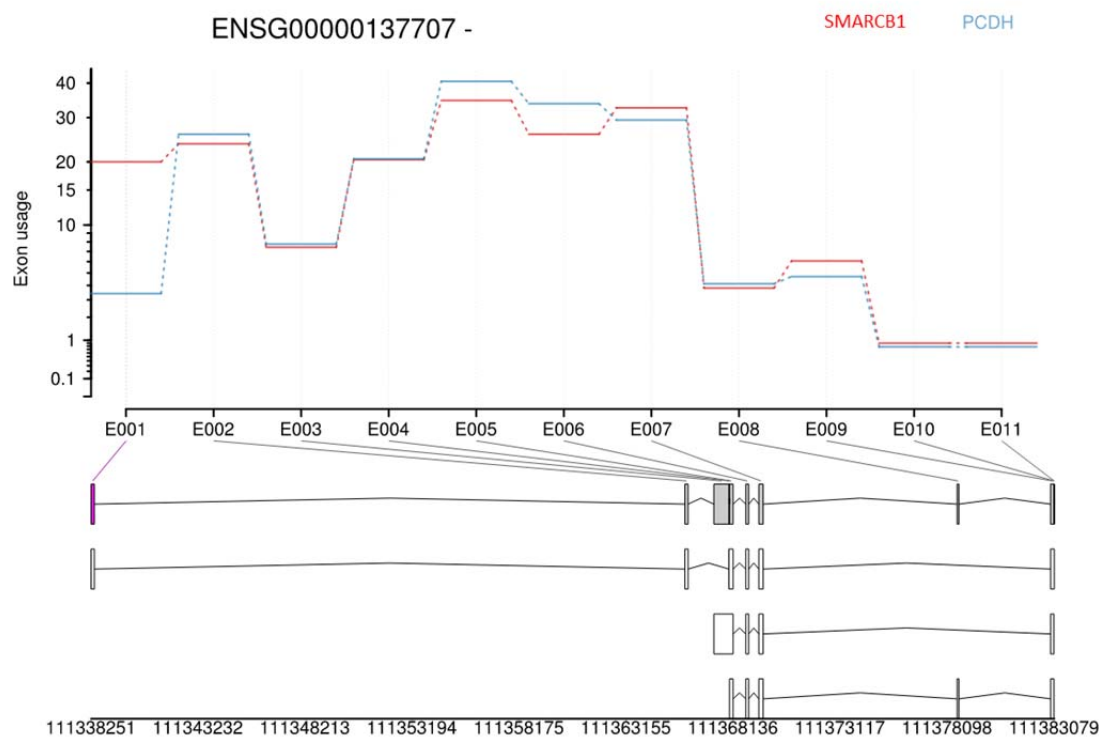


Figure 6.13 *BTG4* shows differential exon usage in response to *SMARCB1* re-expression (contigs 1). Exon usage for *SMARCB1* re-expressing Rhabdoid cell lines are indicated in red (*SMARCB1*) and blue for control infectants (*PCDH*). An annotated transcript model is included at the bottom.

The protein encoded by *PGAM1* plays an important role in cell proliferation; interestingly it has been reported to be up-regulated in many cancers and it also associated with activation of tumour suppressor protein Nm23-H1 (Engel, Seifert et al. 1998). Figure 6.14 shows the differential splicing at an exon level of *BTG4* obtained from DEXSeq analysis. *SMARCB1* re-expression in Rhabdoid cells directly deregulates usage of Exonic parts 4 (exon number 3), activating an alternative splicing of 160 amino acids.

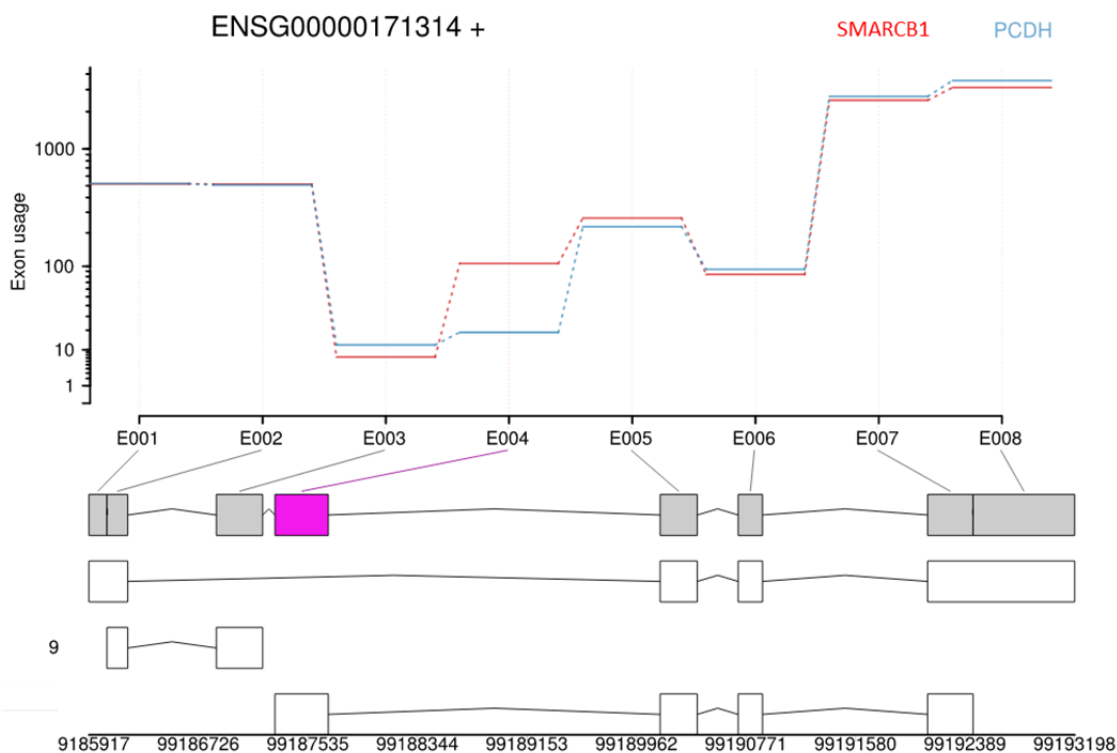


Figure 6.14 *PGAM1* shows differential exon usage in response to *SMARCB1* re-expression (Exonic parts 4). Exon usage for *SMARCB1* re-expressing Rhabdoid cell line is showed in red (*SMARCB1*) and blue for control infectants (*PCDH*). An annotated transcript model is included at the bottom.

#### 6.3.1.6 Validation of the DEXseq analysis

My *SMARCB1* investigation used results retrieved from DEXseq package without any additional filtering, and gave different results from the analysis presented in this chapter. However, this approach lead to false positive results that were either unconvincing when examined by eye or difficult to further validate by RT-PCR i.e. a higher proportion of false positives: exonic parts with low raw counts emerged as significant, mainly because of the absence of raw counts in other exons in the same gene. In the analysis presented here only genes with a high number of normalised reads (>100) in a high number of samples (>75%) were used and this has helped to characterise exons that are truly significant differentially used.

However from the 100 top genes with differentially exon usage obtained using the preliminary analysis, two genes *MYLK* and *IL17RE* were identified as of interest. *MYLK* was regulated by MTA2 in gastric cancer cells modulates the cytoskeleton and participates in cell invasion (Zhou, Ji et al. 2013). Figure 6.15 shows the differential exon usage of *MYLK* obtained with DEXSeq and in particular the different usage at exon 4 (exonic parts 10 and 16). qPCR was performed in all four cell lines after *SMARCB1* re-expression and in the vector control infectants to validate the differentially used exons identified (Figure 6.16), in particular primers were designed to span the differentially used exons identified in the analysis ; *SMARCB1* re-expressing G401 and A204 showed the most significant changes in usage of exon 4 with an increase of usage of 2.7 and 4.8 fold change respectively; STA-WT1 showed a less significant change in usage of exon 4 upon *SMARCB1* re-expression while CHLA-266 did not show change in exon usage of this gene.

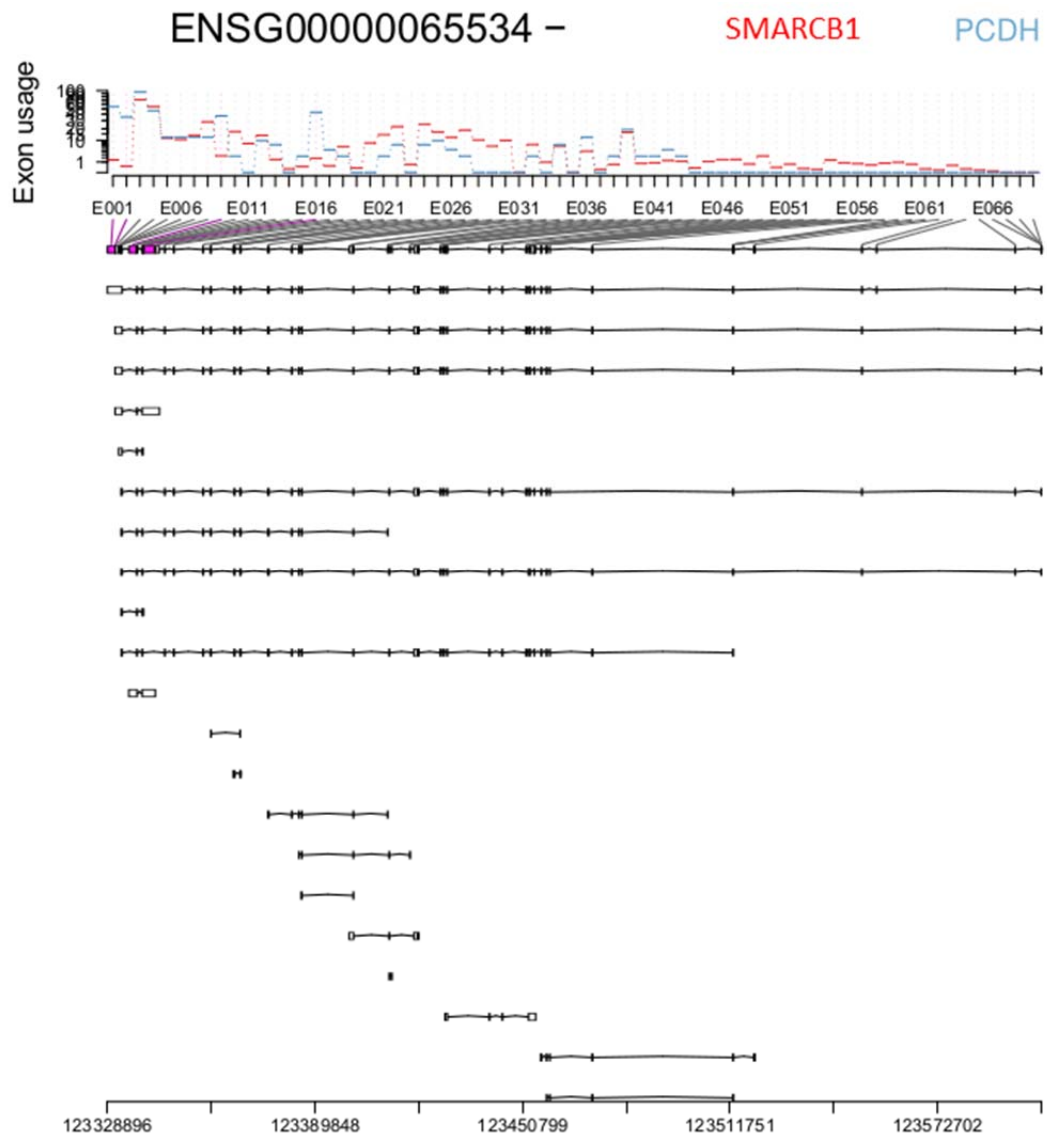
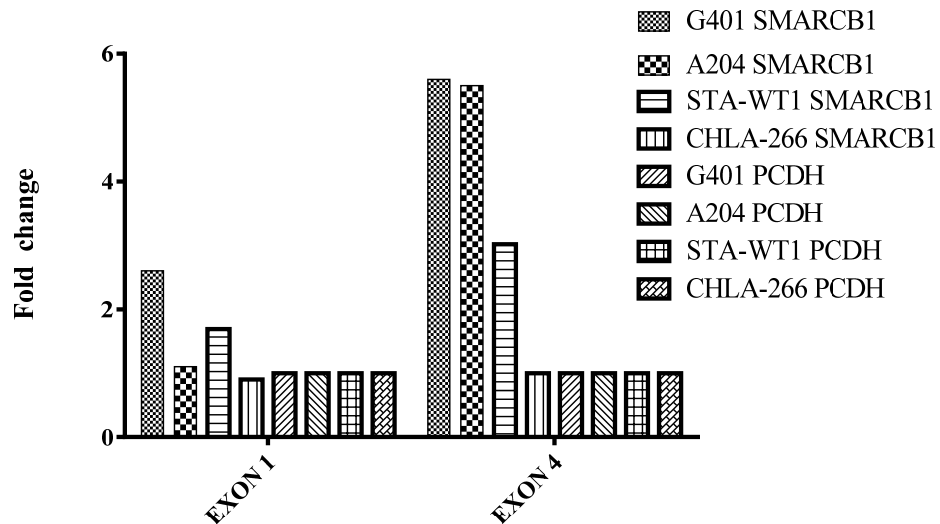


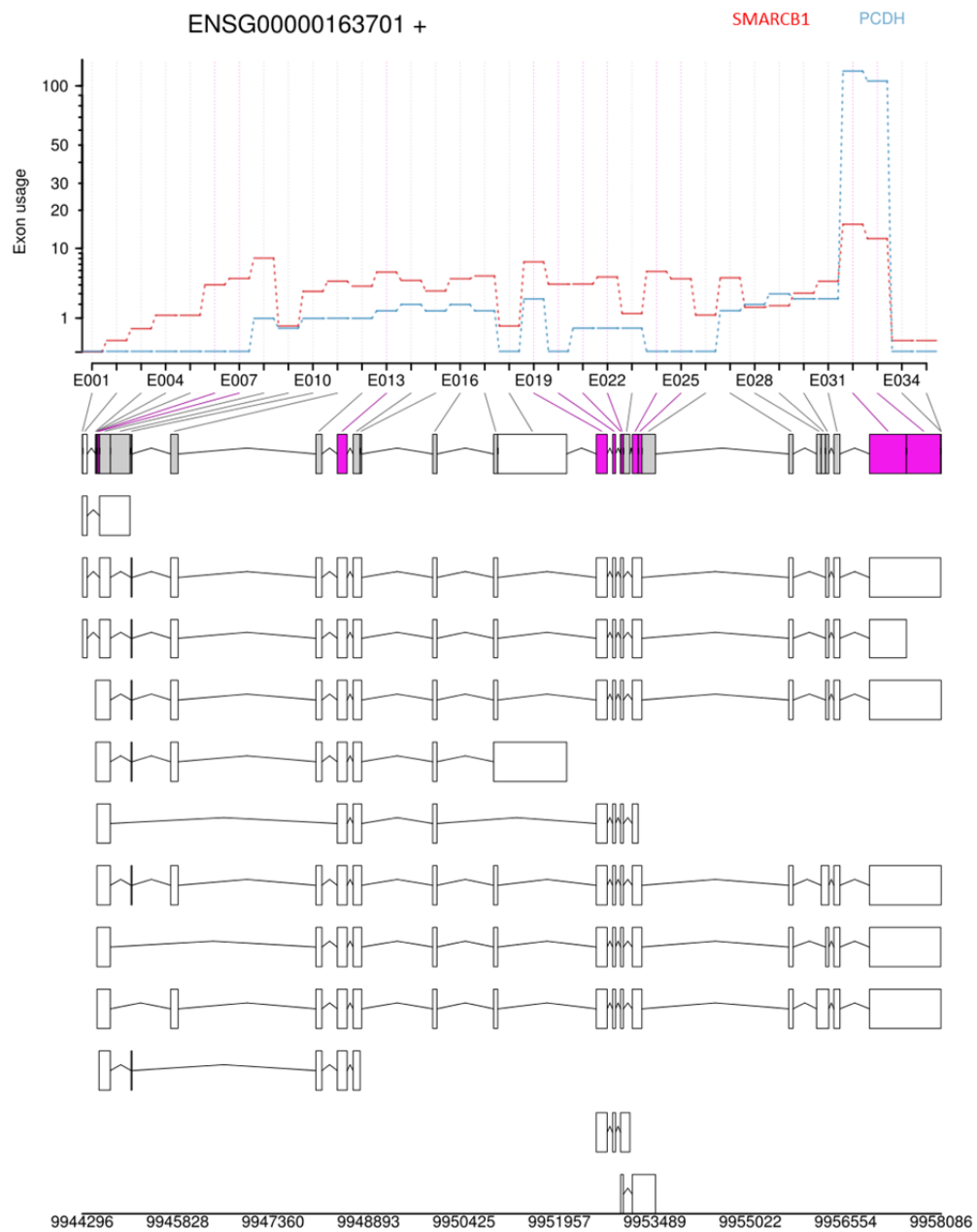
Figure 6.15 MYLK shows differential exon usage in response to SMARCB1 re-expression (Exonic parts 11 and 16). Exon usage for SMARCB1 re-expressing Rhabdoid cell lines are indicated in red (SMARCB1) and control infectants (PCDH) in blue. An annotated transcript model is included at the bottom.



*Figure 6.16 SMARCB1 re-expression induces differential usage of MYLK exon 4 in Malignant Rhabdoid cell lines. RNA was extracted at 7 days post infection with pcdh-mcs1-efl-puro (PCDH) and SMARCB1-pcdh-mcs1-efl-puro (SMARCB1). The mRNA levels were measured by real-time qPCR and normalised for GAPDH. Values are fold change relative to control infectants and are the mean of 3 independent measurements.*

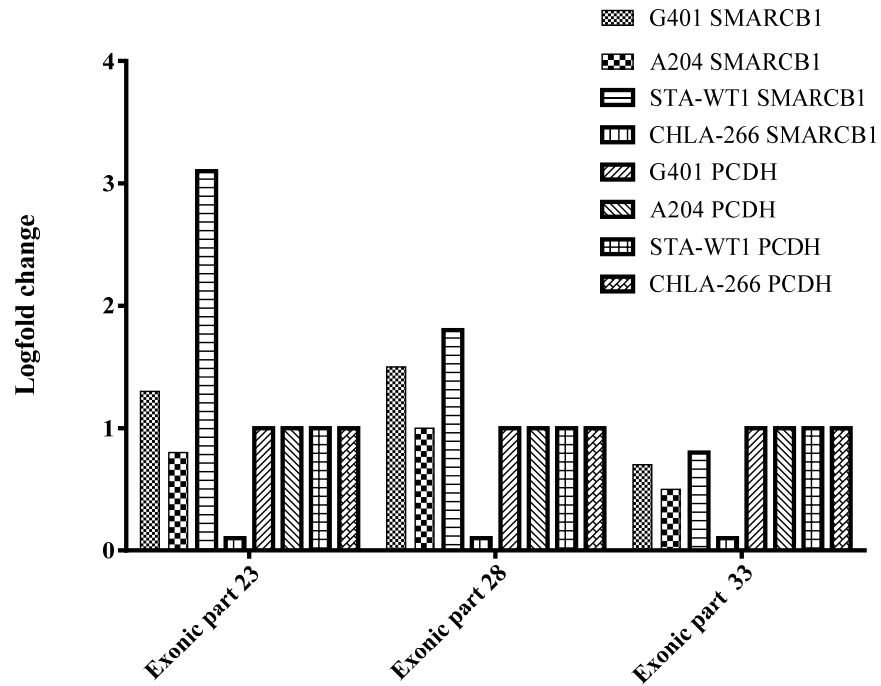
*IL17RE* belong to the cytokines IL-17 and encodes for the IL17RE protein which forms heterodimers with IL17RA; the complex L-17RA-IL-17RE mediates activation of IL-17 signalling in inflammation and participates in tumourigenesis (Al-Samadi, Kouri et al. 2014). Figure 6.17 shows alternative splicing at the exon level for IL17RE; in particular exonic part 33 was significantly differentially used in response to *SMARCB1* re-expression. qPCR was performed in all four cell lines following *SMARCB1* re-expression and in the vector control infectants to validate differentially used exons identified by DEXseq (Figure 6.18); RT-PCRs corresponding to exonic part 23 and 28 were also analysed for experimental comparison as control. Interestingly, STA-WT1 showed a significant decrease of exon usage at exonic part 33 (corresponding to exon 16) with a change in exon usage of -4 fold change when compared with exonic part 23

(exon 10). Decrease of expression was also detected in G401; however CHLA-266 and A204 cell lines did not show a significant change in exon usage of this gene.



*Figure 6.17 IL17RE differential exon usage in response to SMARCB1 re-expression (contigs 5-12). Splicing for SMARCB1 re-expressing Rhabdoid cell line are indicated in red (SMARCB1) and blue for control infectants(PCDH). An annotated transcript model is included at the bottom.*





*Figure 6.18 SMARCB1 re-expression induces differential exon usage of IL17RE exonic part 23, 28 and 33 in Malignant Rhabdoid cell lines. RNA was extracted at 7 days post infection with pcdh-mcs1-ef1-puro (PCDH) and SMARCB1-pcdh-mcs1-ef1-puro (SMARCB1). The mRNA levels were measured by real-time qPCR and normalised for GAPDH. Values are the mean of 3 independent measurements. Exonic parts 23 and 28 were also analysed for experimental comparison.*

### 6.3.1.7 Novel transcript discovery

Following the alignment of the reads to the Human Genome hg19, *De novo* transcript assembly was performed separately on each primary Rhabdoid Tumour sample and each Rhabdoid cell line using Cufflinks, as described in section 3.1.8. The individual *de novo* transcriptomes were merged with one another and with the GENCODEv17 transcriptome library into a single transcriptome using Cuffmerge. Cuffdiff was used to identify significantly differentially expressed genes and splicing isoforms between *SMARCB1* re-expressing Rhabdoid cell lines and control transfected cell lines to generate adjusted p-values, mean FPKM (Fragments Per Kilobases per Million Reads) values and fold change for each gene or isoform.

I focused on characterising genes that do not overlap with the GENCODE 17 transcript library and isolated those unknown genes whose expression was significantly altered by *SMARCB1* expression; data were further filtered to characterise significant gene (minimum FPKM (Fragments per Kilobases per million) in at least one condition > 5). Table 6.3 shows the 12 novel transcripts as result of this filtering.

Gene ID	Position	Log2fold	pvalue
<b>XLOC_008212</b>	chr10:123237667-123491375	-1.04304	0.0013
<b>XLOC_006183</b>	chr10:29656052-29660582	-2.60765	5.00E-05
<b>XLOC_027266</b>	chr18:22096547-22242163	-1.6866	0.001
<b>XLOC_027881</b>	chr18:22242527-22251803	-1.79878	0.0021
<b>XLOC_032441</b>	chr2:150959288-150975090	-1.91077	5.00E-05
<b>XLOC_034832</b>	chr2:189896598-190305848	2.71254	5.00E-05
<b>XLOC_035957</b>	chr20:3019951-3053163	-1.16294	5.00E-05
<b>XLOC_038910</b>	chr22:49599095-49778198	-1.16185	0.0003
<b>XLOC_050753</b>	chr6:122017749-122409661	-1.58329	5.00E-05
<b>XLOC_053533</b>	chr7:40165570-41173105	3.01944	5.00E-05
<b>XLOC_057406</b>	chr8:129733656-129931429	-1.38871	5.00E-05

Table 6.3 List of significant possible novel transcript individualised in *de novo* transcriptome analysis. hg19 position in the genome is included.

These novel transcripts were visualised using a custom track in UCSC Genome Browser. In particular, the track corresponding to the prediction program GenScan (<http://genes.mit.edu/GENSCAN.html>) was used to find any potential predicted genes within the putative de novo transcriptome sequences, and finally the tracks were evaluated against other UCSC annotations. For instance, information such as absence of long non-coding RNA (lincRNA), H3K27Ac marks close to active regulatory elements, DNase hypersensitivity and conservation amongst species (Table 6.4) were considered. Read alignments retrieved from *SMARCB1* re-expressing cells and control infectants for those transcripts identified as potential novel transcript were compared using Integrative

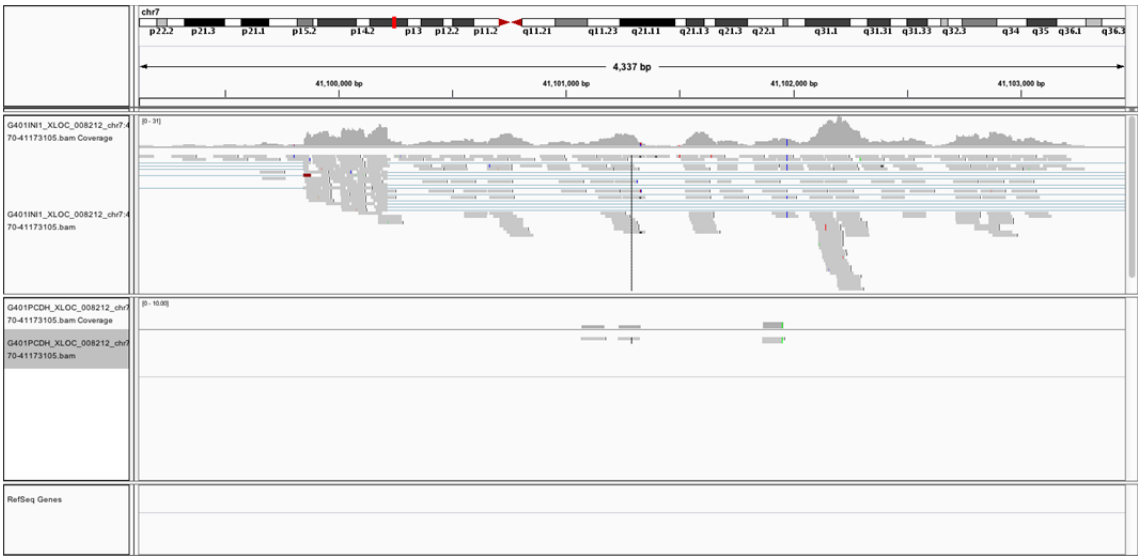
Gene ID	GenScan prediction	High conservation	DNA Hyper- sensitivity	H3K27Ac marks
<b>XLOC_008212</b>	Yes	Yes	Yes	Yes
<b>XLOC_006183</b>	Yes	Yes	No	No
<b>XLOC_027266</b>	Yes	Yes	Yes	Yes
<b>XLOC_027881</b>	Yes	Yes	No	No
<b>XLOC_032441</b>	Yes	Yes	Yes	Yes
<b>XLOC_034832</b>	Yes	Yes	No	Yes
<b>XLOC_035957</b>	Yes	Yes	No	No
<b>XLOC_038910</b>	Yes	Yes	Yes	No
<b>XLOC_050753</b>	Yes	Yes	Yes	No
<b>XLOC_053533</b>	Yes	Yes	Yes	Yes
<b>XLOC_057406</b>	Yes	Yes	Yes	No

*Table 6.4 Analysis of significant transcript characterised by de novo transcriptome analysis. For each transcript functional annotations are indicated.*

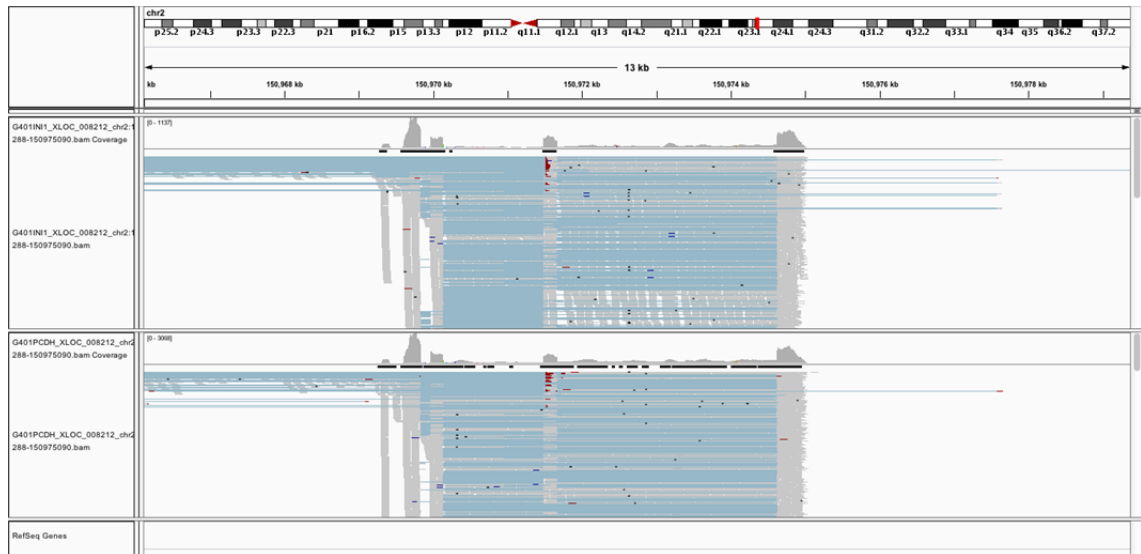
From 12 potentially novel transcripts 2 new potentially transcriptional fragments (XLOC\_008212, XLOC\_053533) and a new potential gene (XLOC\_032441) were identified. In this study XLOC\_053533 and XLOC\_032441 were further analysed to validate the presence of novel transcripts.

### Transcript validation

The read alignment for these two novel transcripts was verified using Integrative Genomics Viewer (Broad Institute). Alignment of reads was essential to identify exon boundaries, For instance, 5 exons were found for XLOC\_053533 (Figure 6.20), while 6 exons were determined for XLOC\_032441 (Figure 6.20).



*Figure 6.19 IGV representation of raw read counts that correspond to XLOC\_053533 and the difference between SMARCB1 re-expressing cells and the control infectants.*



*Figure 6.20 IGV representation of raw read counts that correspond to and XLOC\_032441 and the difference between SMARCB1 re-expressing cells and the control infectants. XLOC\_032441 presents reads in both conditions; however the read count number in SMARCB1 re-expressing cells were significantly higher.*

The exon boundaries identified for XLOC\_053533 were Initially validated by RT-PCR in all four Rhabdoid cell lines, using as a control D283, D425 and DAOY Medulloblastoma cell lines. (Figure 6.21 and Figure 6.22). However, *SMARCB1* dependency appeared most evident for exon boundary 3-4 and 4-5 and especially in the cell line A204.

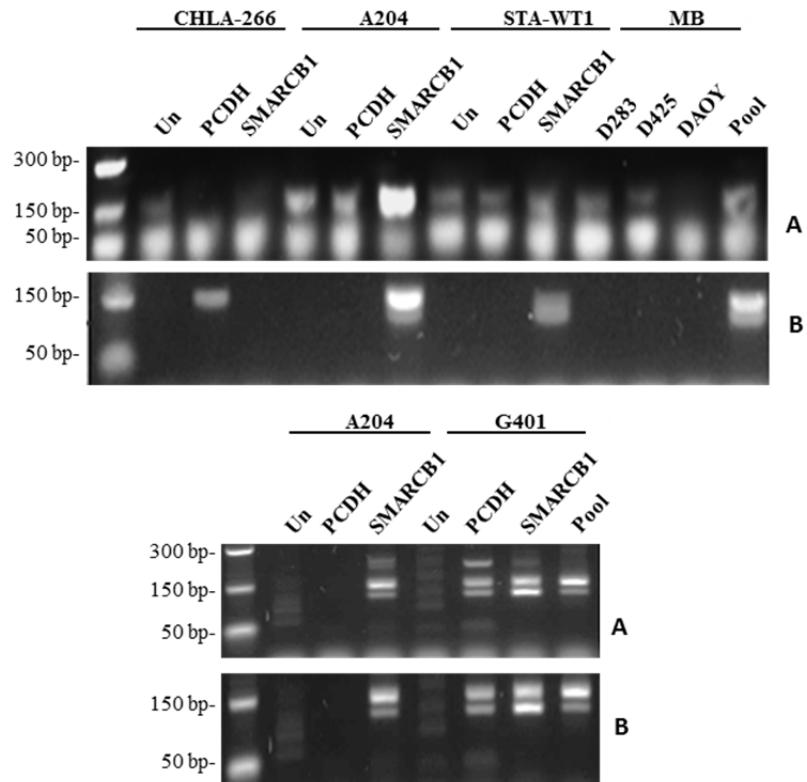
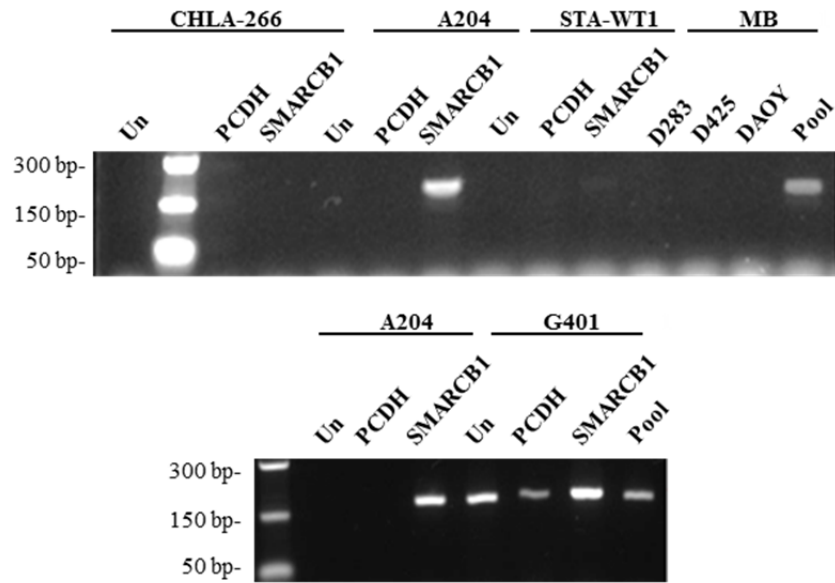
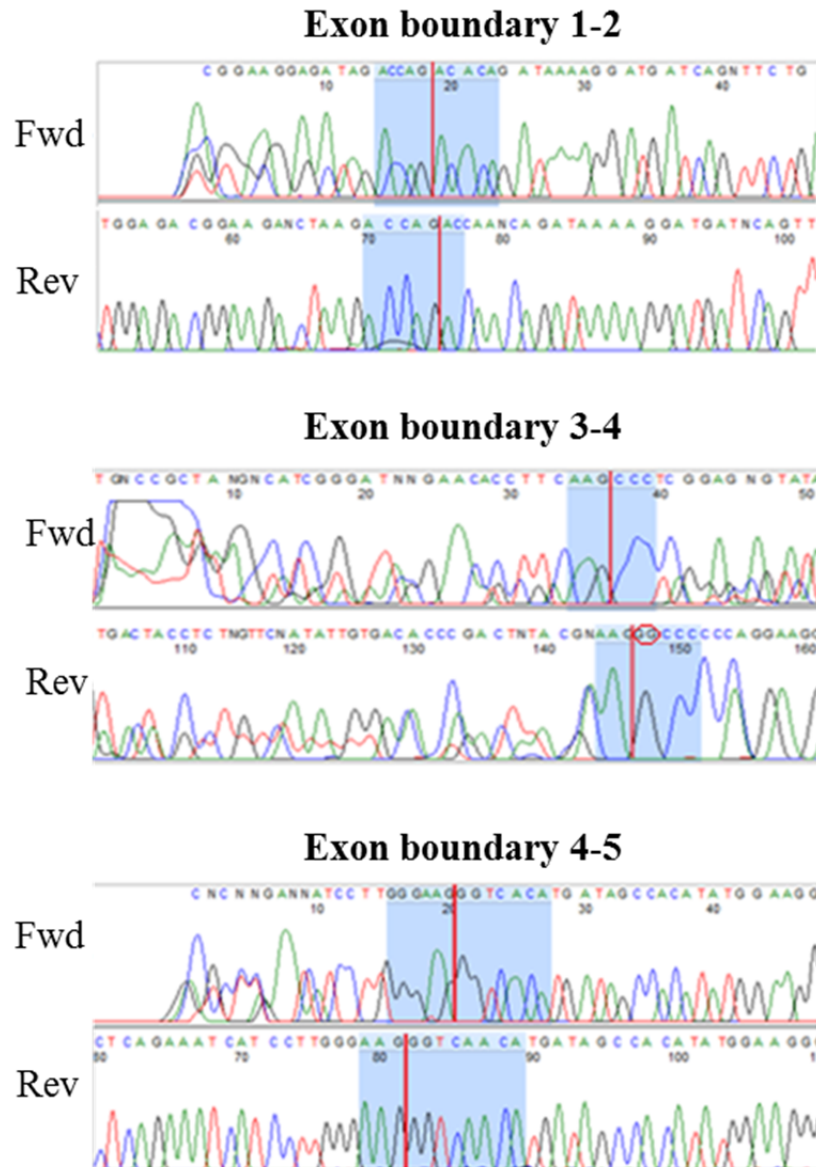


Figure 6.21 RT-PCR for *XLOC\_053533* exon boundary 1- 2(A) and exon boundary 3-4(B). RT-PCR was performed for each Rhabdoid cell line in three different condition: untreated cells (Un), vector control infectants (PCDH) and *SMARCB1* re-expressing cells (*SMARCB1*; also a pool of *SMARCB1* re-expressing cells was used. As a control three Medulloblastoma cell lines (MB) were used.



*Figure 6.22 RT-PCR for XLOC\_053533 exon boundary 4-5. RT-PCR was performed for each Rhabdoid cell line in three different conditions: untreated cells (Un), vector control infectants (PCDH) and SMARCB1 re-expressing cells (SMARCB1); also a pool of SMARCB1 re-expressing cells was used. As a control three Medulloblastoma cell lines were used.*

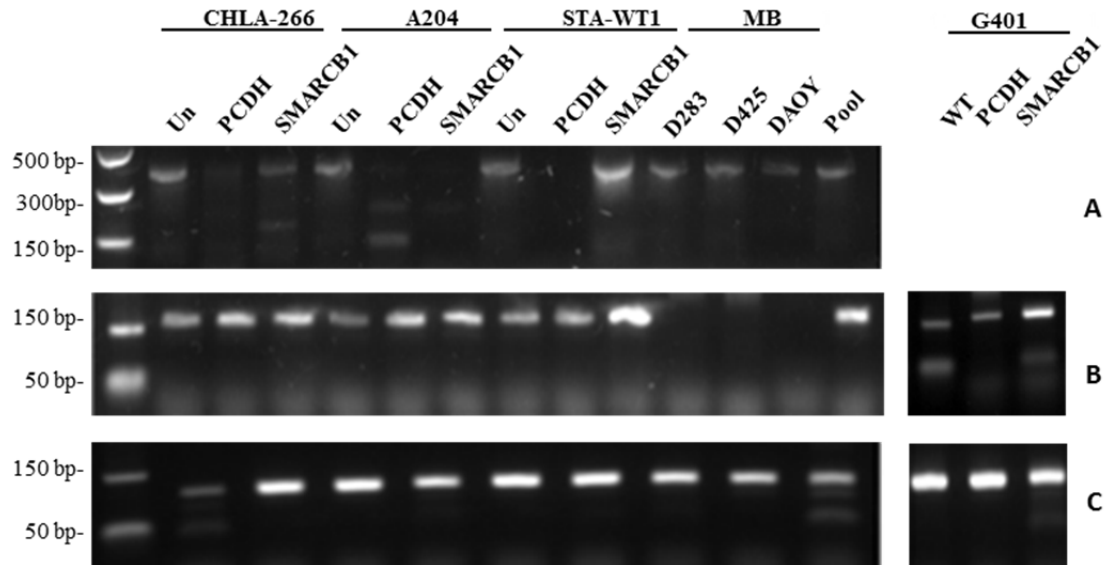
RT-PCR products obtained from *SMARCB1* re-expressing pool cells were sequenced to verify respectively PCR products spanning exons 1 and 2, spanning exon 3 and 4, spanning exons 4 and 5. Sanger sequencing confirmed the presence of the exons boundaries (Figure 6.23).



*Figure 6.23 Sanger sequencing of XLOC\_053533 validated the presence of exon boundary 1-2, exon boundary 3-4, and exon boundary 4-5 at the locus chr7:40165570-41173105.*



For the novel transcript XLOC\_032441, five of the six exons identified were validated by RT-PCR (Figure 6.24), which confirmed the presence of an unknown gene at this locus on Chromosome 2. However the dependence of expression on *SMARCB1* cannot be established using RT-PCR although may be evident by quantitative RT-PCR.



*Figure 6.24 RT-PCR for XLOC\_032441 exon boundary 1-2 (A), exon boundary 2-3(B) and exon boundary 4-5 (C). Except for exon boundary 1- 2 RT-PCR was performed for each Rhabdoid cell line in three different conditions: untreated cells (Un), vector control infectants (PCDH) and SMARCB1 re-expressing cells (SMARCB1); also a pool of SMARCB1 re-expressing cells was used. As a control three Medulloblastoma cell lines (MB) were used.*

RT-PCR products obtained from *SMARCB1* re-expressing pool cells were sequenced to verify respectively spanning exons 2 and 3. Sanger sequencing confirmed the presence of this exon boundary only in the forward direction (Figure 6.25). Other boundaries could not be validated at this time.



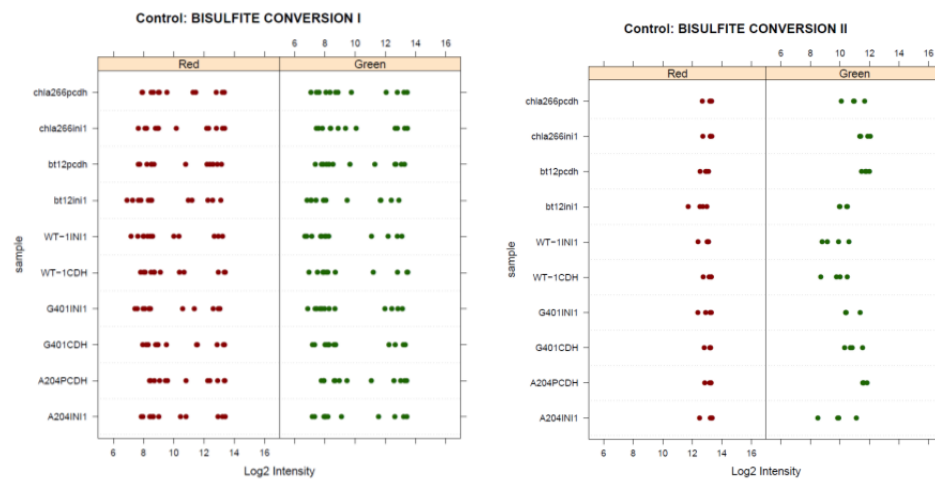
*Figure 6.25 Sanger sequencing validated the presence of exon boundary 2-3 in the locus XLOC\_032441 chr2:150959288-150975090.*

The obtained sequences were further analysed using the homology research tool BLAST. Interestingly, XLOC\_053533 sequences showed homology with *PTRTJ* gene (Chr11: 48101421-48101441) also the XLOC\_032441 sequence showed homology with genes *ILIRAPL1* (ChrX: 28910337-28915794) and *PKHDI* (Chr6: 51803515-51809).

## 6.3.2 450K methylation analysis

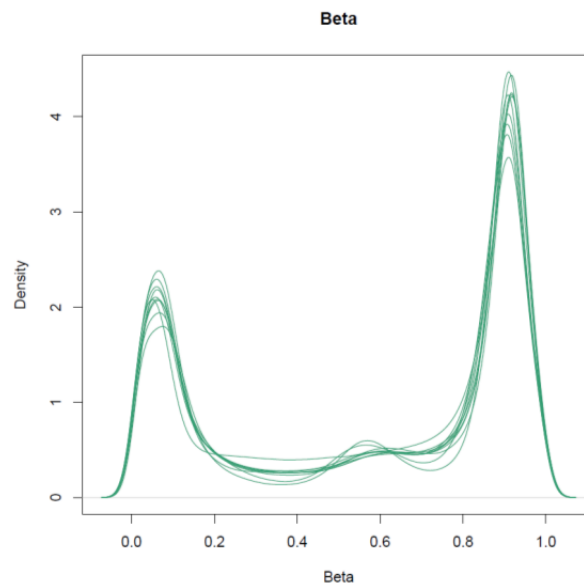
### 6.3.2.1 QC analysis

Quality of the data was performed using Minfi package, as described in section 3.3.2. As a result of the pre-processing step the methylated and unmethylated control probe signals were evaluated for both InfI (Infinium I) and InfII (Infinium II) chemistries (Figure 6.26).



*Figure 6.26 Control strip plot of 450K methylation data for Rhabdoid cell line models. Quality evaluation for InfI (Infinium I A) and InfII (Infinium II B) is indicated. Distribution of signal intensity for each of the quality control probes in both is plotted. Control probes were successfully converted given that samples present roughly similar distribution in both channels.*

$\beta$  values were derived using the BASH algorithm (section 3.2.2): and their distribution analysed to determine the quality of the data. All of the cell line samples pass the quality control checks with less than one percent of CpG sites possessing detection p-values equal or greater than 0.001 (Figure 6.27).



*Figure 6.27 Overall density distributions of beta values in Rhabdoid cell line models. Kernel density plot illustrates the bimodal distribution of  $\beta$ -values for the 43 primary tumours analysed.*

#### *6.3.2.2 Global patterns of DNA methylation in Rhabdoid Tumours*

The distribution of average methylation status within test cohorts (9506 probes selected as those CpGs were most significantly different when *SMARCB1* was re-expressed;  $p < 0.05$ ) is summarised in Table 6.5. Interestingly, a very marked difference in the

methylation status of probes located within and outside of CpG islands was detected: majority of the probes were located outside the CpG Island and unmethylated.

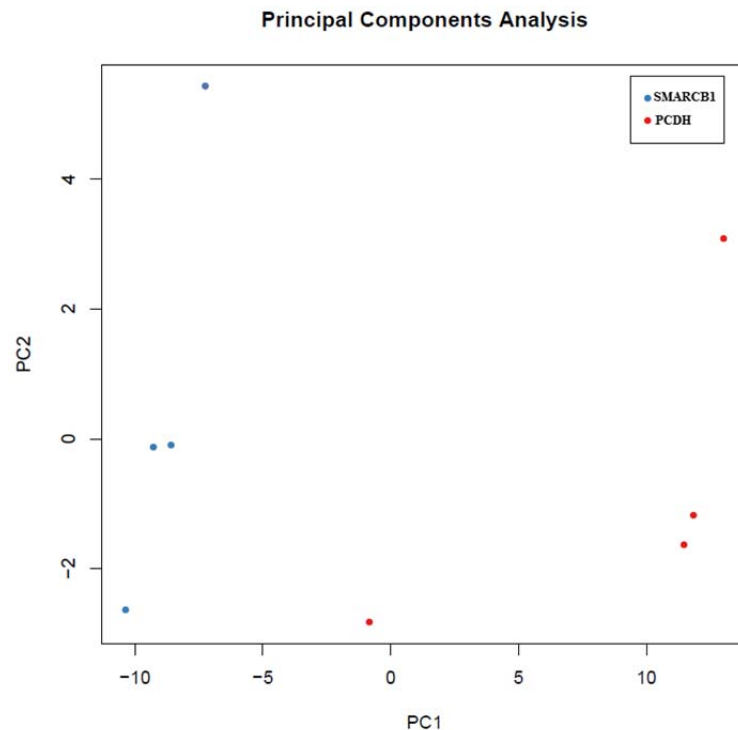
	<b>Probes classification</b>	<b>%</b>
<b>All probes</b>	Methylated	3803 (40%)
	Hemimethylated	86 (0.9%)
	Unmethylated	5615 (59.1%)
<b>CpG Island</b>	Methylated	2597 (27.6%)
	Hemimethylated	1 (0.1%)
	Unmethylated	1181 (12.4%)
<b>Non CpG Island</b>	Methylated	1208 (12.7%)
	Hemimethylated	85 (0.9%)
	Unmethylated	4434 (46.6%)

*Table 6.5 Global methylation distribution among the cell lines (9506 probes selected as those most variable probes within the data). Average methylation scores were calculated for all probes across the datasets. For three probe categories (all probes, probes located within CpG islands and probes located outside of CpG islands), the number of unmethylated*

### *Cluster analysis*

450K methylation data were filtered in a supervised fashion to create a subset composed of the most significantly differentially methylated CpGs (assessed from 40000 most variable CPG probes subsequently tested by limma to an adjusted pvalue <0.01). PCA analysis was performed using a subset of 9506 probes and confirms a genome-wide change in methylation status related to *SMARCB1* re-expression (Figure 6.28). It is also notable that one of the cell lines does not alter CHLA266 does not alter its methylation

status to the same degree upon *SMARCB1* re-expression. Changes in methylation are generally smaller in this line and it is uncertain at this point why.



*Figure 6.28 Supervised principal component analysis on SMARCB1 re-expressing cell lines vs control infectants. 450K methylation data distinguishes two groups. Red points represent vector control infected cells (PCDH) while SMARCB1 re-expressing cells (SMARCB1) are indicated in blue. The x-axis is the first component; the y-axis is the second principal component.*

Consensus NMF supervised analysis (k=2) was carried out on the selected genes identifying two different metagenes one for the *SMARCB1* positive cells and one for the control cells (Figure 6.29). As can be seen, one of the cell lines CHLA266 continues to group preferentially with the control lines even when *SMARCB1* is re-expressed. An examination of the individual metagene scores shows that whilst the shift in the

*SMARCB1* dependent methylation metagene is equivalent in the CHLA266 line is far more methylated at the relevant CpG loci to begin with and therefore the shift in methylation provoked by *SMARCB1* is insufficient to cause the consensus NMF to group this cell line with the other *SMARCB1* re-expressing lines. Nevertheless, the same CpGs are being altered and in the same direction in the CHLA266 so there is no need to remove from the analysis or to alter my results.

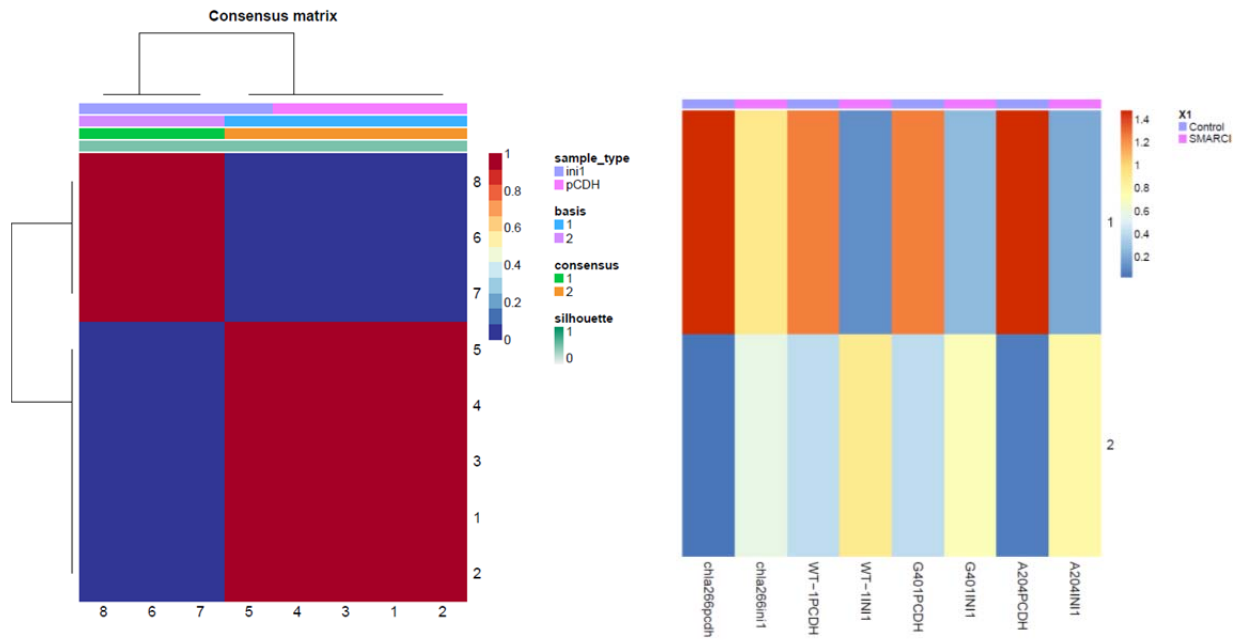


Figure 6.29 (left) Supervised consensus NMF and hierarchical cluster analysis ( $K=2$ ) performed on methylation data from *SMARCB1* re-expressing cells (*SMARCB1*) and control infectants (*pCDH*). The matrices were coloured 0 (deep blue, indicating samples never in the same cluster) to 1 (dark red, indicating samples always in the same cluster). (right) a heatmap of the methylation metagenes  $k=2$  extracted from *SMARCB1* re-expressing cells (*SMARCB1*) and control infectants (*PCDH*) methylation profiles. Shading shows the shift in metagene expression following *SMARCB1* expression is equivalent in direction but not of equivalent magnitude in the cell line CHLA266 (first two columns)

Filtered data were used to generate a heatmap (Figure 6.30) revealing that *SMARCB1* re-expression causes up regulation of the majority of the genes analysed again it can be seen that the magnitude of hypo and hypermethylation in CHLA266 following *SMARCB1* expression (1<sup>st</sup> column) is not equivalent in magnitude.

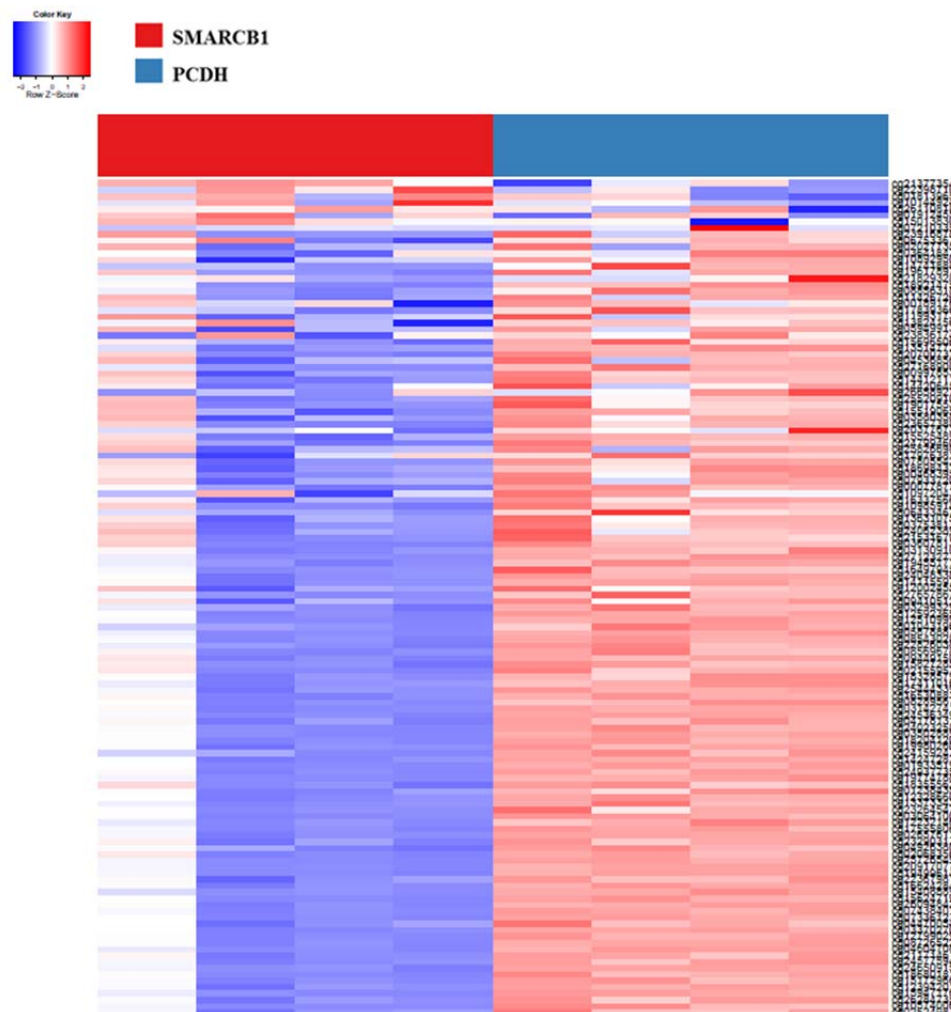


Figure 6.30 Supervised heatmap of differentially methylated probes in *SMARCB1* re-expressing Rhabdoid cells (*SMARCB1*) and in vector control infected Rhabdoid cells (*PCDH*). Overrepresented (pale red to dark red) and underrepresented transcripts (pale blue to dark blue). Each row represents a



probes and each column a sample. Overrepresented (pale red to dark red) and underrepresented transcripts (pale blue to dark blue).

Table 6.6 shows the 15 most differentially methylated CpGs as a result of filtering. Eight of the gene linked to these CpGs were associated with tumour progression (\*), two promoting metastasis (+), one indirectly involved in tumour invasion and angiogenesis (-), one gene indicated as oncogene (~) (<http://www.genecrads.org>). Complete list in Appendix I.

Probes	Gene Symbol	Average Delta values	LogFold change	Adjusted p value
cg20377673	<b>MIF*</b>	-0.63678	-3.92831	0.021081
cg02524983	<b>LPP*+</b>	-0.62799	-3.90218	0.00116
cg18680181	<b>KIAA0391</b>	-0.62626	-3.84263	0.009082
cg24650915	<b>SIL1*</b>	-0.62301	-3.86597	0.019342
cg21171461	<b>RAPGEF1</b>	-0.62141	-4.01879	0.019342
cg08726522	<b>ST5~</b>	-0.61973	-4.62047	0.021081
cg12799029	<b>ABR</b>	-0.60408	-3.92576	0.020375
cg26094842	<b>VTI1A</b>	-0.6016	-4.06983	0.015605
cg04604708	<b>CREB3L2*</b>	-0.5998	-3.84498	0.002931
cg20917077	<b>CPA5</b>	-0.59935	-4.17312	0.019342
cg15947176	<b>C8orf46*</b>	0.47743	5.201754	0.013222
cg24577594	<b>GALNT2</b>	0.871935	3.677821	0.021081
cg07438401	<b>SUCLG2</b>	-0.63678	-3.92831	0.009082
cg21377354	<b>SUSD2*-+</b>	-0.62799	-3.90218	0.000312
cg22398710	<b>FBXL5*</b>	-0.62626	-3.84263	0.037258

*Table 6.6 Selected differentially methylated CpGs in response to SMARCB1 re-expression. LogFold Change is given as a ratio of M-values between SMARCB1 re-expressing Rhabdoid cells and vector control infected Rhabdoid cells. Average delta values are the mean change in beta value between SMARCB1 re-expressing Rhabdoid cells and vector control infected Rhabdoid cells Functional annotation of genes relevant to cancer is shown in the table with \* for genes associated with tumor progression, + for genes promoting metastasis, - for genes correlated with tumour<sub>3</sub> invasion and angiogenesis, ~ for genes indicated as oncogenes.*

Interestingly *SMARCB1* re-expression induces hypomethylation of *ST5*; gene included also in the top 20 most differentially methylated genes when Extra Cranial Rhabdoid malignancies were compared with Atypical Teratoid Rhabdoid Tumours. *ST5* encodes for p126 protein which interacts with MAP kinase activation pathways in response to EGF (Majidi, Hubbs et al. 1998). A smaller isoform p70 is also encoded which is considered a tumour suppressor gene (Majidi, Hubbs et al. 1998), however the mechanism is not yet understood. One of the two hypermethylated genes is *SUSD2*, which has been recently linked with breast tumourigenesis; the encoded protein interacts with Gal-1 and the interaction promotes tumour immune evasion, angiogenesis, and metastasis(Watson, Evans et al. 2013).

## **6.4 RNA-seq analysis identifies gene reactivation induced by 5-aza-2-deoxycytidine treatment in Malignant Rhabdoid cells**

RNA-seq was carried out in the four Rhabdoid cell lines following treatment with the demethylating agent 5-aza-2-deoxycytidine (5-azaCdR) to try to elucidate the genome-wide effect on the epigenetic mechanisms involved in Rhabdoid tumourigenesis.

### **6.4.1 RNA-seq analysis**

#### *6.4.1.1 Primary samples preparation*

RNA from treated cells was extracted as described previously in section 2.1.1. RNA integrity number (RIN score) was evaluated for each sample, using an Agilent Bioanalyzer. All the RNA samples were scored with a RIN greater than 9.

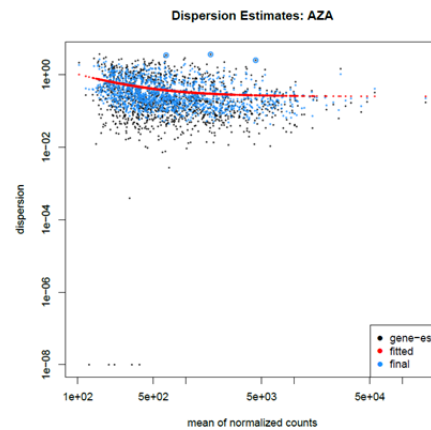
#### *6.4.1.2 RNA-seq QC analysis*

Samples sequenced successfully passed the RNA quality control checks: all bases analyses showed an average Phred-score higher than 30. High variability on the first bases as result of random priming was also found, confirming that this effect was not related with sample preparation or perform of sequencing.

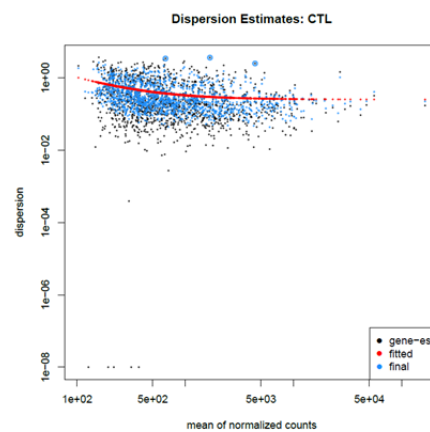
#### *6.4.1.3 DEseq analysis*

Raw count values were analysed in a pair-wise fashion using DESeq2 in R. Differentially expressed genes between cells treated with 5-azaCdR and untreated cells

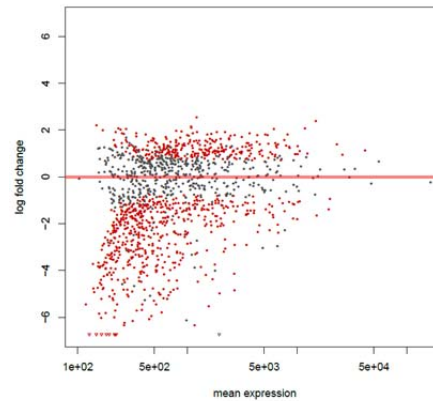
were determined by the negative binomial method of statistical analysis. Dispersion in the data set was estimated (Figure 6.31 and Figure 6.32), as well as differential expression estimation between cell lines with or without 5-azaCdR treatment (Figure 6.33).



*Figure 6.31 Gene expression dispersion estimation and model fitting of 5-azaCdR treated cells (AZA). Each dot represents per gene dispersion estimation and the red line the fitted mean dispersion function.*



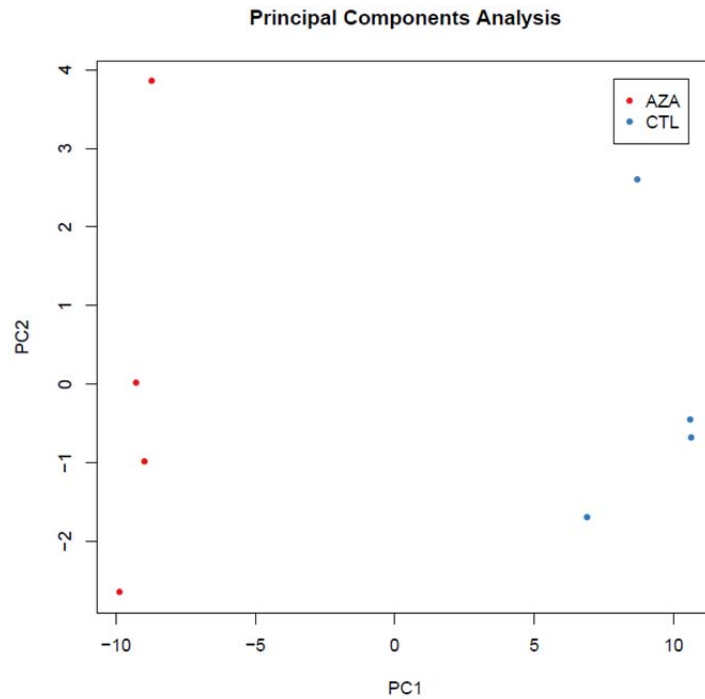
*Figure 6.32 Gene expression dispersion estimation and model fitting of control untreated cell lines. Each dot represents per gene dispersion estimation and the red line the fitted mean dispersion function.*



*Figure 6.33 Testing for differential expression between of 5-azaCdR treated cells vs untreated cells: differentially expressed genes Log2fold versus mean expression over all samples. Red dots represent genes marked as differentially expressed at 10% FDR, when a Benjamini and Hochberg method is applied. The lines on the bottom of the plot represent genes with a very low or very high log fold change (value off scale).*

### *Clustering analysis*

A subset of significant differentially expressed gene was extracted from the DEseq results to identify genes significantly differentially expressed in response to 5azaCdR treatment (adjust  $p$ -value  $< 0.01$ ; moderated log fold change  $\leq -2$  and  $\geq +2$ ). PCA analysis was performed on a subset of 132 significant genes, and identifies two separate subgroups, one consisting of untreated cells and the second of 5-azaCdR treated cells (Figure 6.34). This result demonstrates the concerted effect on gene expression of demethylation induced by 5-azaCdR.



*Figure 6.34 Supervised principal component analysis of gene expression profiles from 5-azaCdR treated Rhabdoid cell lines and untreated Rhabdoid cell lines distinguishes two different biological groups. Points in red represent 5-azaCdR treated cells (AZA) and in blue vector control untreated cells (CTL). The x-axis is the first component; the y-axis is the second principal component.*

Consensus NMF supervised analysis (k=2) was carried out on the selected genes identifying two different metagenes, and confirming the change in expression caused by 5-azaCdR (Figure 6.35).

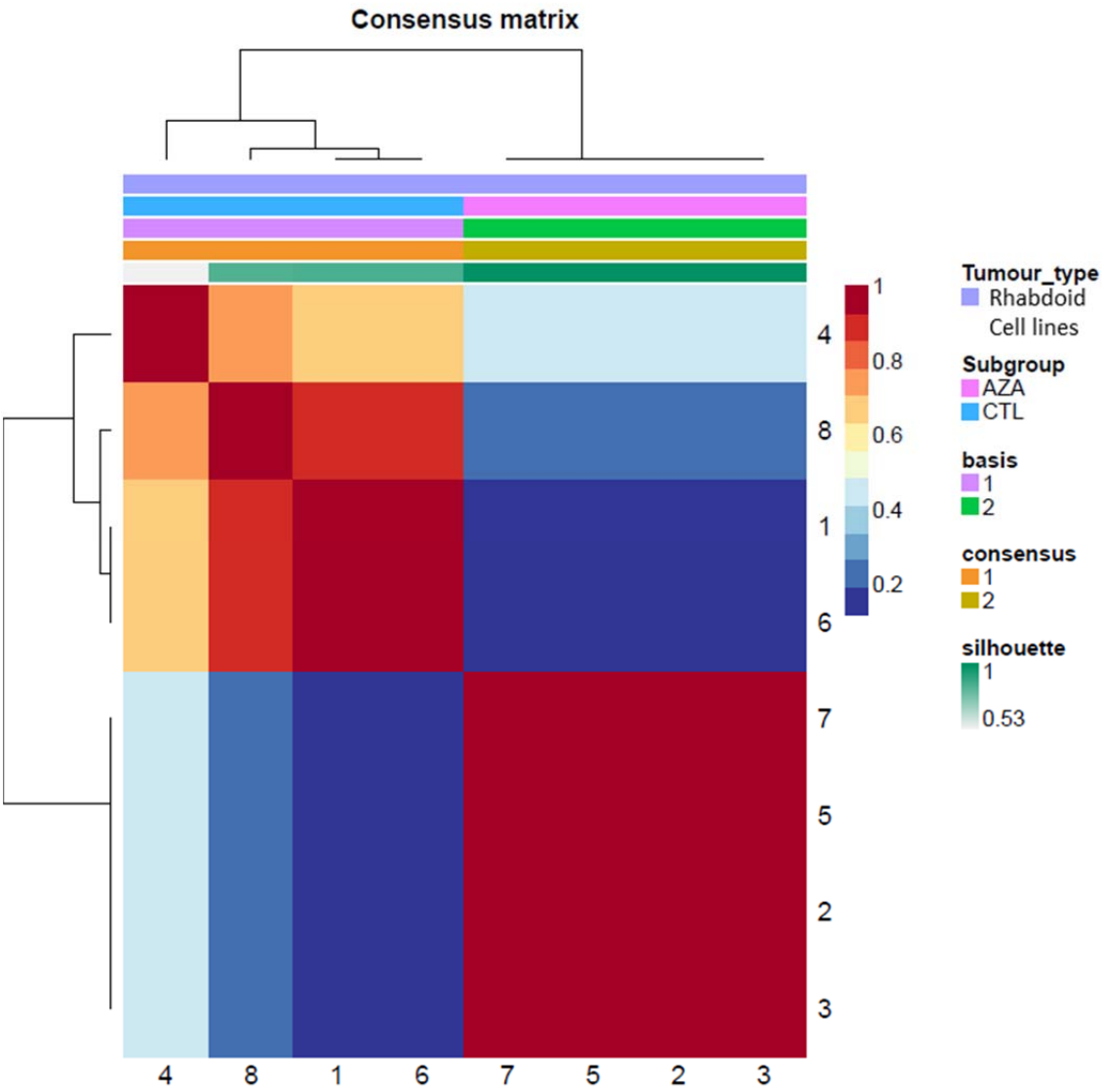


Figure 6.35 Supervised NMF and hierarchical cluster analysis (K=2) of 5-azaCdR treated cells (AZA) and untreated cells (CTL) DEseq data. The

matrices are coloured 0 (deep blue, indicating samples never in the same cluster) to 1 (dark red, indicating samples always in the same cluster).

These significant differentially expressed filtered genes were used to generate a heatmap revealing that 5-azaCdR treatment causes up regulation of the majority of the genes analysed (Figure 6.36).

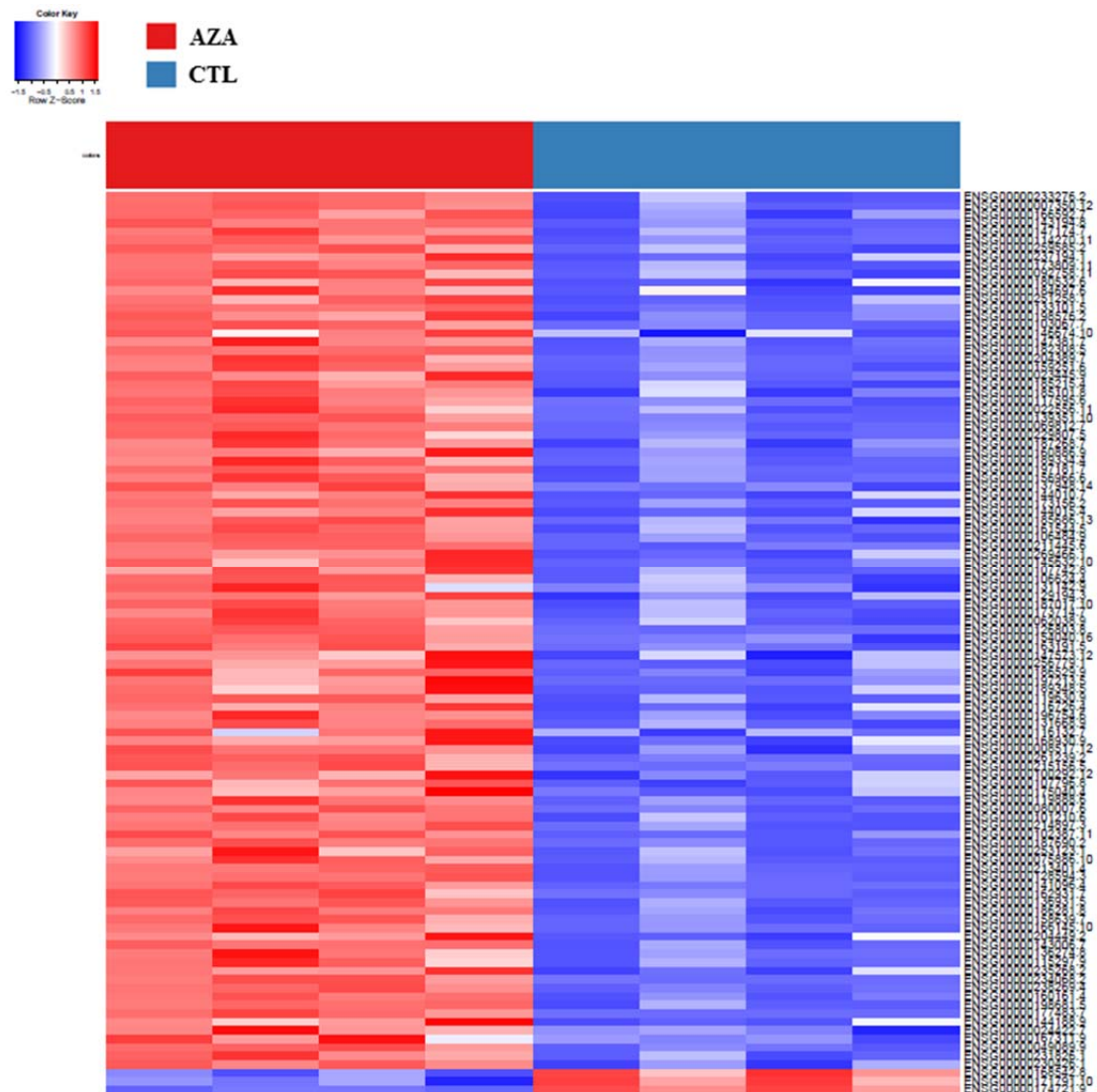


Figure 6.36 Supervised heatmap of selected significantly differentially expressed genes in 5-azaCdR treated cells (AZA) and untreated cells (CTL).



*Overexpressed (pale red to dark red) and under-expressed transcripts (pale blue to dark blue). Each row represents a gene and each column a sample.*

Table 6.7 shows the 15 most differentially expressed genes ranked by fold change. Seven genes are associated with tumour progression (\*) and one gene indicated as oncogene (~) (<http://www.genecrads.org>). Complete list in Appendix I.

Gene Symbol	Adjusted <i>p</i> -value	Moderate log2fold
<b>GPX1 *</b>	1.45E-06	4.207726
<b>TKTL1*</b>	2.63E-07	4.114281
<b>RRAD*</b>	2.93E-09	4.097312
<b>MAEL</b>	3.25E-15	3.804456
<b>ACRC</b>	1.97E-10	3.417875
<b>COL7A1*</b>	7.36E-26	3.402389
<b>TDRD12</b>	3.64E-08	3.133676
<b>COL9A3</b>	8.77E-10	3.049116
<b>ZSCAN4 *</b>	1.61E-05	2.993179
<b>CLDN6*</b>	0.000668	2.954359
<b>RFPL4B</b>	8.23E-10	2.942049
<b>CCNA1*</b>	2.09E-20	2.887184
<b>COL3A1</b>	2.01E-08	-2.00126
<b>BCL2 ~</b>	1.47E-07	-2.00128
<b>FGFBP3</b>	7.52E-17	-2.31185

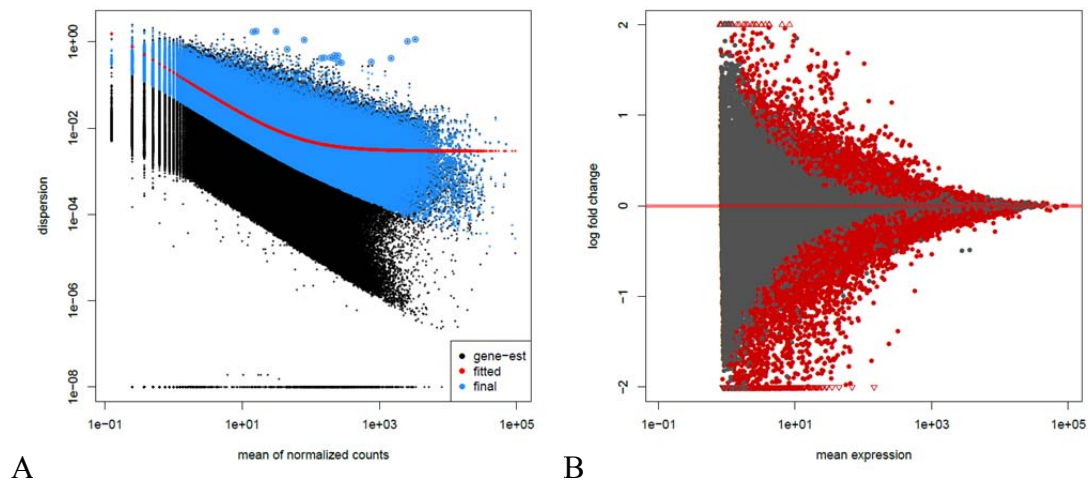
*Table 6.7 Differentially expressed gene SMARCB1 re-expressing Rhabdoid cells vs vector control infected Rhabdoid cells. Functional annotation of genes relevant to cancer is shown in the table with \* for genes associated with tumor progression, ~ for genes indicated as oncogenes.*

*GPX1* encodes for the anonymous antioxidant proteins, which plays an important role in inflammatory response; in *GPX1* together with *GPX2* knockout mice Polycomb (PcG) target genes show changed in methylation status especially in embryonic stem cells (Hahn, Hahn et al. 2008). *ZSCAN4* is newly identified embryonic stem cell marker and

recently a direct interaction with Rap1 was described; interestingly *Rap1* causes senescence, regulating the expression of genes encoding for core histone proteins; downregulation of *Rap1* increases core histone protein levels causing global changes in chromatin stability and therefore in gene expression (Lee and Gollahon 2014).

#### *DEXSeq analysis: differential exon usage*

Alternative splicing and alternative isoform expression were identified from row count values using a negative binomial method. Dispersion estimate in the data and distribution of exon usage for each gene were also computed comparing untreated cells and 5-azaCdR treated cells (Figure 6.37).



*Figure 6.37 A) Per exon usage dispersion estimation and model fitting (A) and testing for differential exon usage of genes (B) comparing untreated cells and 5-azaCdR treated cells. A) Each dot represents per exon dispersion estimation and the red line the fitted mean dispersion function. B) Log2fold versus mean expression over all samples. Red dots represent exons marked as differentially expressed at 10% FDR, when Benjamini and Hochberg method is applied. The lines on the top and bottom of the plot*

*represent exons with a very low or very high log fold change (value off scale).*

#### *6.4.1.4 Detection of differentially used exon following by 5-aza-2-deoxycytidine treatment*

The exon level counts were examined by comparing using DEXseq 5-aza-CdR treated cells with untreated cells in order to determine a methylation dependent effect upon exon usage. The DEXseq results were filtered for values with an adjusted *p*-value greater than 0.05 and log2 fold change less than -1 or greater than 1. Moreover, only protein coding genes were analysed; 7196 genes with significant variation in exon usage were identified. Table 6.8 shows the 15 most differentially used exons as ranked by log fold change. From the 15 genes in the table, five were associated with tumour progression (\*), one gene is indicated as an oncogene (~) (<http://www.genecrads.org>). Complete list in Appendix I.

*PCAM* represents an interesting gene, since its role is somewhat ambiguous: a high level of expression denotes an aggressive phenotype of oesophageal cancer and it is used as biomarker to characterise circulating tumour cells, however absence or low expression of *EPCAM* in circulating tumour cells is also reported (Driemel, Kremling et al. 2014). *PDE4D* encodes a cAMP-specific phosphodiesterases protein and interestingly cAMP levels are inversely correlated with tumour progression and grade of malignancy in brain (Sengupta, Sun et al. 2011).

Gene Symbol	Exonic part	Adjust <i>p</i> value	log2fold
<b>BCAR3 ~</b>	E030	6.52E-09	2.951094
<b>EPCAM*</b>	E002	3.28E-08	2.714227
<b>BBS9</b>	E004	1.43E-10	2.500092
<b>CPEB4*</b>	E012	4.49E-10	2.359447
<b>RAD51B</b>	E055	0.000303	2.134738
<b>TENM4</b>	E046	4.66E-05	2.129385
<b>PDE4D*</b>	E037	0.011333	1.96927
<b>WDR43</b>	E002	0.001071	-0.79376
<b>RPSAP52</b>	E002	0.008397	-0.80263
<b>AC003090.1</b>	E007	0.001825	-0.8259
<b>PSMG3</b>	E007	1.56E-08	-0.83298
<b>RP11-10A14.4</b>	E004	0.007052	-0.88819
<b>CHCHD4*</b>	E006	1.63E-15	-0.8913
<b>WRAP53*</b>	E005	3.56E-05	-0.91958
<b>TNIP2</b>	E012	0.00252	-0.98382

*Table 6.8 Genes with differentially used exons in 5-azaCdR treated Rhabdoid cells and untreated cells. Functional annotation of genes relevant to cancer is shown in the table with \* for genes associated with tumour progression, ~ for oncogenes.*

## 6.5 Discussion

### 6.5.1 Integrated genomic and epigenetic analysis identifies downstream target genes dependent by *SMARCB1*

Earlier studies have indicated that re-expression of *SMARCB1* in Rhabdoid cell lines causes broad upregulation of the expression of genes involved in cell cycle, cell growth and cytoskeleton organization (Medjkane, Novikov et al. 2004, Morozov, Lee et al. 2007). Only one study attempted to characterise the downstream effectors of *SMARCB1*, however the analysis was limited to one cell line (Morozov, Lee et al. 2007).

RNA-seq was performed in Malignant Rhabdoid Tumour cells upon *SMARCB1* re-expression and accomplished in four cell lines derived from different location of the body. Gene expression analysis reveals that *SMARCB1* re-expression significantly altered 493 genes and 471 were detected to be upregulated. Interestingly, *DNA2*, *SERPINF1* and *LBX1* were in the top 15 deregulated genes. Deregulation of *DNA2* causes DNA instability as result of chromosomal aberrations and inter-nuclear chromatin bridges. In Malignant Rhabdoid Tumours *SERPINF1* was already identified as a potential therapeutic target (Gadd, Sredni et al. 2010). *LBX1* is a gene involved in neuro-development and interestingly high expression level is associated with aggressive phenotype in breast cancer. Exon usage analysis of Rhabdoid cell lines indicated that re-expression of *SMARCB1* directly affects the exon usage of 493 genes. Many of the top 100 genes were linked with tumour progression and neurodevelopment; in this list *FAM60A*, *BTG4*, *PGAM1*, *MYLK* and *IL17RE* were identified as potential therapeutic targets; *FAM60A* interacts to chromatin remodel E2F1, which is not usually expressed in Malignant Rhabdoid Tumours (Venneti, Le et al. 2011); *BTG4* is expressed in embryos and down-regulation of this gene promotes tumour cell growth in gastric cancer (Dong, Tu et al. 2009); *PGAM1* also is involved in cell proliferation and in particular associated with activation of tumour suppressor protein Nm23-H1 (Engel, Seifert et al. 1998) *MYLK* modulates the cytoskeleton and participates in cell invasion (Zhou, Ji et al.

2013). The protein encoded by *IL17RE* forms a complex with IL17RA, which mediates inflammatory signalling in the brain (Al-Samadi, Kouri et al. 2014). Differential exon usage of *MYLK* and *IL17RE* was also analysed by qPCR confirming the different splicing observed through DEXseq analysis; however only in part as CHLA-266 and A204 cell lines did not show a significant change in exon usage of *IL17RE* gene.

*De novo* transcriptome analysis was also performed on RNA-seq raw data to identify novel transcripts and new isoforms. Comparison of the *de novo* transcriptome with GENCODE17 transcript library identified 12 possible novel transcripts with expression dependent on *SMARCB1* re-expression; further analysis identified one possible new isoform and a novel gene: XLOC\_053533 and XLOC\_032441. Exon boundaries and relative structures were validated by RT-PCR and Sanger sequencing; BLAST analysis of the obtained sequences confirmed the presence of the isoform and the gene; however both sequences showed homology for other genes and in particular XLOC\_053533 for lincRNA.

Also in all four cell lines after *SMARCB1* re-expression, analysis of the methylation status was performed; 450K methylation array was carried out in Rhabdoid cell line to study the epigenetic effects dependent of *SMARCB1* loss. Our analysis indicated that *SMARCB1* re-expression causes mainly hypomethylation of several genes, in particular in this analysis 61 significantly hypomethylated probes and just 2 significantly hypermethylated. The top 15 most differentially methylated genes included *ST5* specifically hypomethylation following *SMARCB1* re-expression; interestingly this gene was also significantly differentially methylated in the comparative analysis between Extra Cranial and Atypical Teratoid Rhabdoid Tumours. Instead *SUSD2* appeared to be hypermethylated following *SMARCB1* re-expression; this gene is indirectly linked with the activity of Gal-1, promoting angiogenesis, and metastasis (Watson, Evans et al. 2013).

The results obtained in this analysis implicate *SMARCB1* involvement in the epigenetic mechanisms that regulate gene expression. In particular, we here show the direct effect of *SMARCB1* loss in gene expression and upon methylation. As a result of this

investigation we observed that *SMARCB1* loss causes deregulation of gene not only involved in cell cycle (as previously described), but also those directly involved in tumourigenesis, metastasis or others with unknown oncogenic activity. 450K array analysis suggests that gene expression in Malignant Rhabdoid Tumours is caused in part by aberrant methylation: broad upregulation of genes was observed concomitant with a general hypomethylation. Interestingly *SMARCB1* re-expression induces mainly hypomethylation of non-CpG island-associated promoters and therefore deregulation of gene expression; the effect of hypomethylation in non-CpG island genes is not a fully understood mechanism.

*SMARCB1* loss not only affects gene expression by changing methylation status, but also activates different isoforms: in particular *SMARCB1* re-expression induces differentially exon usage on 59 genes and therefore expression of different transcript. Most of the genes with differentially exon usage upon *SMARCB1* re-expression are related to tumourigenesis; for example *FAM60A* which promotes tumourigenesis by interaction with chromatin remodelling E2F1 and *BTG4* with its downregulation and hypermethylation is related to tumour proliferation.

In conclusion, this integrated analysis characterises the role of *SMARCB1* on regulation of epigenetic events in Rhabdoid tumourigenesis. In particular, we show that deletion of *SMARCB1* does not just affect cell cycle controlling genes, but a variety of genes as results of a wide epigenetic deregulation, affecting methylation status and alternative splicing.

### **6.5.2 RNA-seq reveals gene activation in Malignant Rhabdoid cell lines following 5-Aza-Deoxycytidine treatment**

5-azaCdR treatment in Malignant Rhabdoid cell lines causes deregulation of genes which are known to be *SMARCB1* dependent (see Chapter 4). To obtain a comprehensive understanding of gene deregulation in Rhabdoid Tumours and the

relationship with methylation status, genome-wide RNA sequencing in four Rhabdoid cell lines after 5-azaCdR treatment was performed.

Study of differential expression identified a total of 132 significant deregulated genes: 116 were upregulated and 16 downregulated in response to 5-azaCdR treatment. Most of the genes were related with tumourigenesis and in particular downregulation of the oncogene *BCL2* was observed, together with upregulation of *GPXI* and *ZSCAN4*. *GPXI* and *ZSCAN4* are both involved in chromatin stability in stem cells, the first modulates the methylation of the Polycomb (PcG) target genes while the second regulates together with Rap1 core histone protein expression (Lee and Gollahon 2014).

5-azaCdR treatment in Rhabdoid tumour cells also affects the expression of alternative isoforms, In particular, we identified 7196 gene with significant alternative isoform regulation in response to 5-aza-CdR treatment. Once again the most significant genes of interest present with a tumourigenic potential or function, such as *EPCAM* and *PDE4D*. High expression of *EPCAM* is observed in aggressive cancer stem cells, while *PDE4D* is reported linked with tumour progression in brain malignancies (Sengupta, Sun et al. 2011, Driemel, Kremling et al. 2014).

This study showed that 5-azaCdR treatment in Rhabdoid tumour cell lines has a broad effect on the genome, in a similar manner to that of *SMARCB1* re-expression. In particular, a general upregulation of gene expression was observed as well as deregulation of splicing control. This confirms the direct involvement of *SMARCB1* and SWI/SNF complexes in genome-wide methylation, resulting in gene expression deregulation. Since aberrant DNA methylation drives tumourigenesis in Malignant Rhabdoid Tumours, inhibitors of methylation such as 5-azaCdR might conceivably be used as a therapy to combat Malignant Rhabdoid Tumours tumourigenesis.



**Chapter 7      Comparison of Molecular  
Signatures from Malignant  
Rhabdoid Tumours and Cell line  
Models Identifies Mechanisms of  
Tumourigenesis**

## 7.1 Introduction

Whole-genome approaches are used to identify the molecular profiles which may be used to identify and define diseases. Whole genome profiling is used to profile gene expression and to study epigenetic modification, illustrating signatures of gene expression and patterns of epigenetic regulation. In Malignant Rhabdoid Tumours whole genome profiling identified prognostic genes and gene profiles that might cause tumourigenesis (Birks, Donson et al. 2013). However, no study has fully characterised epigenetic mechanisms involved in Rhabdoid Tumours. Here integration of data sets from multiple whole genome profiling was attempted, to characterise the epigenetic modulation of DNA methylation that regulate gene expression. Cross-referencing functional models with primary data allows a multi-dimensional characterisation of biological systems, improving the understanding of the contribution of genetic and epigenetic alterations in tumour progression.

Whole genome profiling technology provides the opportunity to identify disease-specific genes and pathways in unprecedented detail. In the past decades studies of disease were mainly based on single gene analysis, an approach that may miss important effects on pathways; however the introduction of high-throughput technologies such as microarray leads to a more comprehensive knowledge of diseases by extracting meaning from a much larger number of differentially expressed genes. The term pathway refers to collections of genes or proteins that directly or indirectly modulate a well-defined biological task. Decades of molecular biology studies have defined a wide number of pathways, establishing functional associations between gene expression and disease. In particular, following the sequencing of the entire human genome accomplished in 2003, pathways analysis has grown exponentially discovering a vast number of significant biological signatures, collected in public pathway repositories (Ramanan, Shen et al.). Although pathway analysis of expression data is a relatively new field, many studies have identified deregulated pathways that reflect perturbed genes from expression data, allowing a better understand of biology of disease. Pathway analysis in the context of Atypical Teratoid Rhabdoid Tumours was previously attempted (Birks, Donson et al. 2011, Knipstein, Birks et al. 2012, Sredni,

Huang et al. 2014). Previous studies underlined the presence in Atypical Teratoid Rhabdoid Tumours of some deregulated pathways and in particular the *BMP* pathway genes as an indicator of poor prognosis (Birks, Donson et al. 2011). Also the involvement of Notch and Wnt signalling pathway as well as IFN pathway has been observed in Rhabdoid tumour cell lines after *SMARCB1* re-expression (Morozov, Lee et al. 2007, Gadd, Sredni et al. 2010).

In this study cross-referencing between RNA-seq and 450K platform was attempted to disclose the relationship between methylation and gene expression. Moreover, in order to understand the mechanism behind *SMARCB1* loss, primary data have been cross-referenced with model data, in an attempt to find gene signatures of Rhabdoid Tumours and gene target that might have a therapeutic relevance. Moreover, a systematic analysis comparing changes in gene expression as a result of *SMARCB1* re-expression and primary tumour profiles was performed to characterise pathways involved in tumorigenesis. In particular, an integrated pathway analysis of the gene expression signature retrieved from RNA-seq metagene analysis was performed to identify downstream *SMARCB1* effectors. Cross-referencing between primary and cell line data was essential to estimate the significance of genes deregulated after *SMARCB1* and the relevance of pathways on Rhabdoid Tumour biology.

## 7.2 Aims

The study reported in this chapter aims to:

- Identify aberrant epigenetic signature dependent on *SMARCB1* loss that cause deregulation of gene expression responsible for cancer development in Rhabdoid Tumours ;
- Identify common significant gene expression signatures by comparing gene expression from Malignant Rhabdoid primary tumour data with *SMARCB1* re-expressing Rhabdoid cell lines;
- Identify pathway(s) key to Rhabdoid tumour development which may ultimately be targetable therapeutically.

### **7.3 Multiple platform data integration characterises the role of *SMARCB1* loss in epigenetic deregulation**

#### **7.3.1 Cross-referencing 450K methylation data primary data and model data**

To assess the role of *SMARCB1* in modulating DNA methylation patterns, methylation data from both Rhabdoid primary tumour samples and *SMARCB1* re-expressing Rhabdoid cell lines were cross-referenced. 450K Rhabdoid primary data sets were analysed to identify significant probes from a subset of 418,698 probes ( $p$ -value  $< 0.01$ ; log Fold Change  $< -3$  and  $> 3$ ). For this study Rhabdoid cell line 450K data sets were re-analysed calculating Beta-values and mean Delta values in a pairwise fashion (e.g. Beta-value of G401 *SMARCB1* – Beta-value of G401 PCDH); probes with a Delta-value  $> 0.2$  and  $< -0.2$  in  $\geq 3$  cell lines were selected. The analysis was focused only on overlapping probes with an observed significant change in methylation in response to *SMARCB1* re-expression.

The significant probes were ranked based upon their Delta-values (the average change in Beta-value) obtained from the primary data. The data were further filtered (Delta-value  $< -0.2$  and log Fold Change  $< 0$ ), to give 63 CpGs consistently hypomethylated in response to *SMARCB1* re-expression and also significantly hypermethylated in the Rhabdoid primary tumour data set as compared to Medulloblastoma. Table 7.1 shows the 20 most significant differentially methylated probes, resulting from the analysis. Complete list in Appendix I.

Probes	Gene Symbol	Average Delta values	LogFold change	Adjusted <i>p</i> value
<b>cg26284735</b>	BHLHE40	-0.552160333	7.413367	9.38E-39
<b>cg24577594</b>	GALNT2	-0.509345351	6.34529	8.72E-43
<b>cg08726522</b>	ST5	-0.503897695	6.302051	1.74E-43
<b>cg12799029</b>	ABR	-0.503414605	6.985098	1.43E-33
<b>cg02657611</b>	TESK2	-0.474253792	6.464976	8.35E-41
<b>cg04926380</b>	EXT2	-0.474071043	4.981123	6.86E-36
<b>cg23264547</b>	EXT2	-0.466024831	5.184367	2.67E-38
<b>cg24159247</b>	ITPR1	-0.454381015	-4.589	2.79E-41
<b>cg16530881</b>	CCR7	-0.44269779	-4.53975	1.22E-31
<b>cg01016122</b>	PSORS1C1	-0.425839909	-4.99831	7.79E-32
<b>cg05238713</b>	ELP4	-0.425812936	-8.27617	6.78E-33
<b>cg01290136</b>	SCFD2	-0.422060232	-7.03077	2.94E-47
<b>cg05033369</b>	FCRLA	-0.416578021	-7.23292	3.08E-42
<b>cg24401262</b>	RP11-439L8.2	-0.401150663	-7.22481	1.67E-51
<b>cg15918685</b>	GOLPH3	-0.398714181	-7.49857	3.37E-47
<b>cg27621721</b>	NMNAT1P3	-0.386671934	-6.90548	2.09E-52
<b>cg00121339</b>	CLEC16A	-0.385790623	-7.49232	3.64E-49
<b>cg03747456</b>	KRT80	-0.383624929	-7.24504	1.71E-51
<b>cg13780614</b>	GALK2	-0.379932013	-7.70371	1.67E-51
<b>cg04384810</b>	C14orf159	-0.379845058	-8.26217	4.13E-48

*Table 7.1 List of top 20 genes and their associated CpG dinucleotides resulting in significant hypomethylation following SMARCB1 re-expression as a result of cross-filtering of primary and model data.*

The top 20 genes were further analysed in order to select those with a function pertaining to cancer and also for those associated with known regulation by chromatin remodelling factors. This analysis was achieved using the UCSC Genome Browser based on Transcription Factor ChIP-Seq data from the Encode project. In particular, the analysis identified seven probes corresponding to *BHLHE40*, *GALNT2*, *ST5*, *ABR*, *CCR7*, *PSORS1C1*, and *KRT80*, based on the presence of these features:

- Evidence according to the ENCODE project of binding sites for SWI/SNF subunits such as SMARCB1, SMARCA4 and/or SMARCA2;
- Relatively high levels of H3K27Ac mark, indicator of active transcriptional region and enhancer sequences (Baylin and Jones 2011);
- Presence of CpG islands located near or at SWI/SNF binding sites;
- Presence of binding sites for POLR2A (subunit of the RNA polymerase II) and CTCF.

#### 7.3.1.1 Validation of selected probes in Rhabdoid cell lines using bisulfite conversion

The selected differentially methylated CpGs were further investigated. Specifically, validation of changes in methylation in these genes was performed by bisulfite PCR in *SMARCB1* re-expressing Rhabdoid cells, followed by Sanger sequencing. Bisulfite PCR to observe the change in methylation status in response to *SMARCB1* expression of three genes was validated: *ABR*, *BHLHE40* and *GALNT2*.

*ABR* regulates members of the Rho family of GTP-binding proteins via a GTPase-activating (GAP) domain and a region with homology to the guanine nucleotide dissociation-stimulating domain (Chuang, Xu et al. 1995). *BHLHE40* is a transcriptional repressor which inhibits proliferation, serum deprivation-induced apoptosis, and activation of pro-caspases associated with the mitochondrial pathway of apoptosis (Wang, Reiser-Erkan et al. 2010). Glycopeptide-preferring polypeptide N-acetylgalatosaminyltransferase (*GALNT1*) initiates O-linked glycosylation of protein by catalysing the attachment of the first N-acetylgalatosamine (GalNAc) residue; knockdown of *GALNT1* reduced proliferation of the bladder cancer cell line and abrogates subcutaneous cancer growth in mice (Gaziel-Sovran, Segura et al. 2011).

# *SMARCB1 re-expression induces changes in ABR methylation status*

CpG dinucleotides of interest were first analysed using the UCSC browser (Figure 7.1). The location of the CpG dinucleotide on the genome is located within  $\approx 3,000$  bp of a CpG island and associated with the promoter of an *ABR* isoform; moreover the regulatory region is marked by high levels of H3K27Ac. Also the presence of binding sites for SMARCB1, SMARCA4, SMARCC1 and SMARCC2 indicate SWI/SNF complex recruitment in this region; POLR2A subunit of RNA polymerase II bind to this region and might induce transcription Initiation in conjunction with the remodelling activity of the SWI/SNF complex.

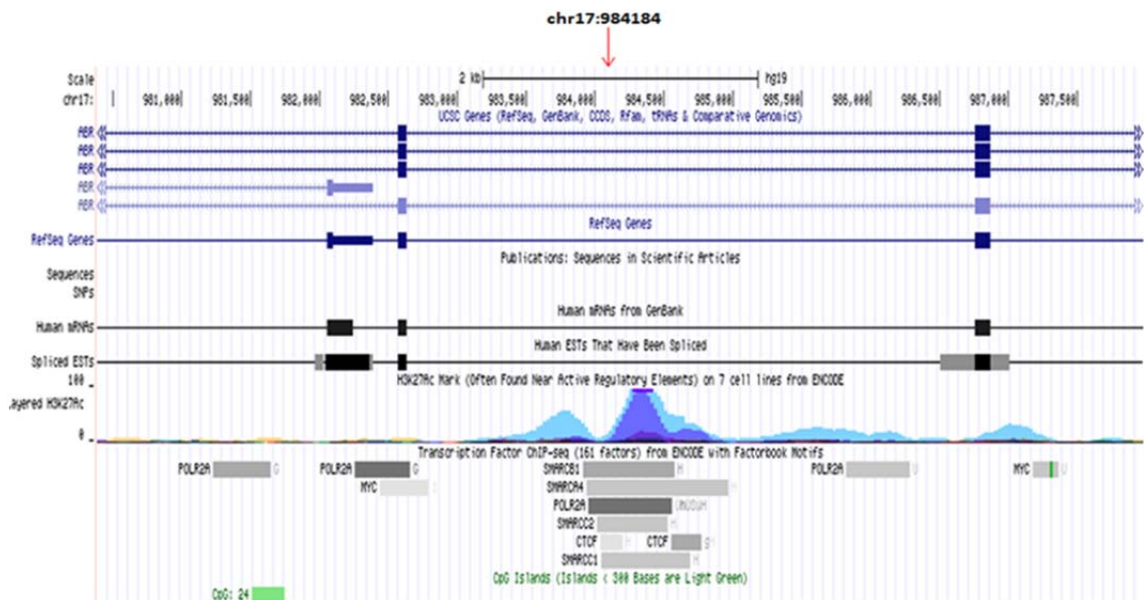


Figure 7.1 Transcription factor ChIP-Seq data for *ABR* derived from the Encode project. The CpG of interest is indicated by the red arrow. Different *ABR* isoforms and Chip-seq annotation from Encode project are indicated.



Sanger sequencing of the bisulfite PCR showed that the CpG dinucleotide at Ch17:9841184 is hypomethylated, as a result of *SMARCB1* re-expression (Figure 7.2). Of the three adjacent CpG dinucleotides identified just one was hypomethylated in all cell lines in response to *SMARCB1* methylation.

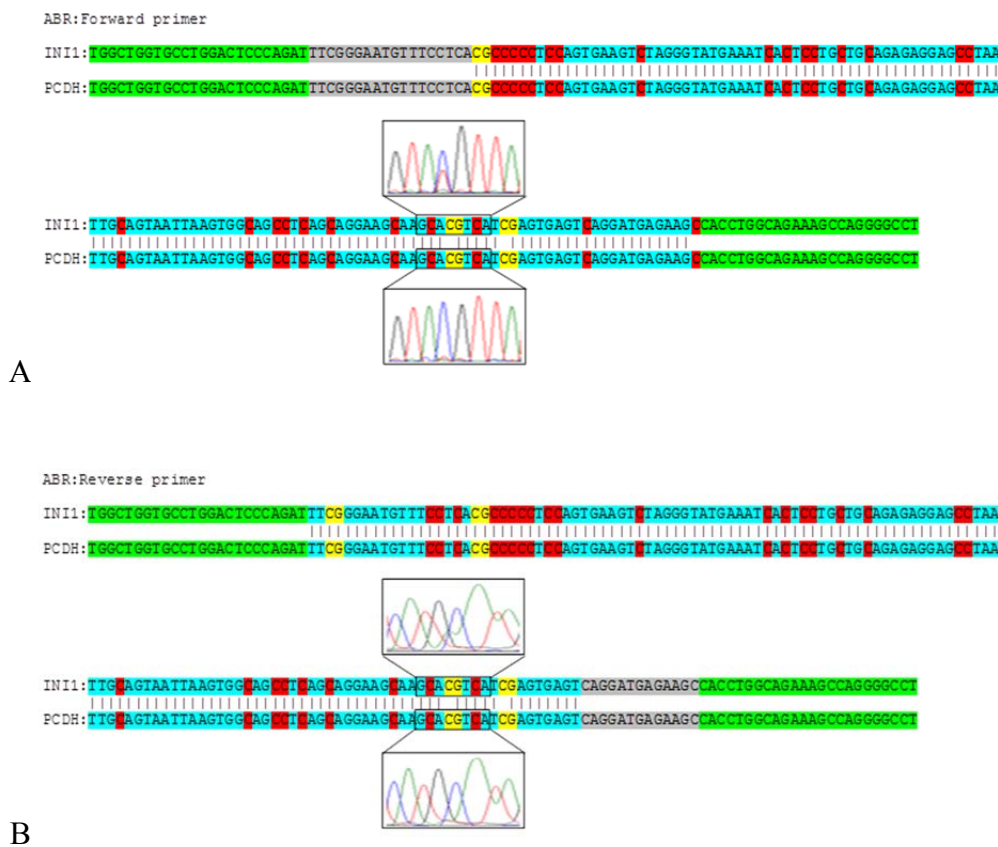
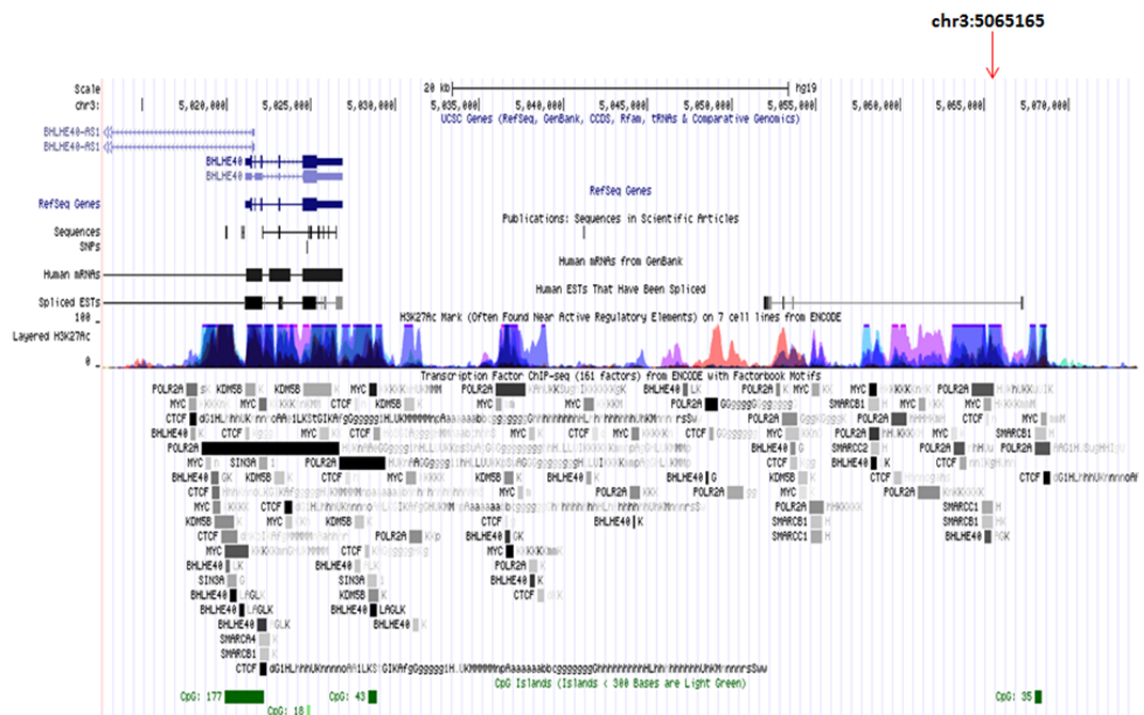


Figure 7.2 Sanger sequencing indicates demethylation of ABR in response to *SMARCB1* re-expression in G401 cell lines. Forward (A) and reverse primers (B) are indicated in green. Blue colour represents matches between *SMARCB1* and *PCDH*. Cytosine bases highlighted in red are unmethylated and converted into Ts and As in the forward and reverse strand, respectively. The trace is given for the CpG of interest (Ch17:9841184). Evidence of demethylation is indicated by overlapping traces.

# *SMARCB1 re-expression induces changes in BHLHE4 methylation status*

The CpG dinucleotide of interest (Chr3:5065165) was first analysed using the UCSC browser. The CpG dinucleotide of interest is located ~35,000 bp downstream of *BHLHE40* within a regulatory region characterised by high levels of the H3K27Ac mark (Figure 7.3 Transcription factor ChIP-Seq data for BHLHE4 derived from the Encode project. The CpG of interest is indicated by the red arrow. Different BHLHE4 Isoforms and Chip-seq annotation from Encode project are indicated.). Binding sites for SMARCB1, SMARCC1, POLR2A and MYC are located in the region proximal to a CpG island.



*Figure 7.3 Transcription factor ChIP-Seq data for BHLHE4 derived from the Encode project. The CpG of interest is indicated by the red arrow. Different BHLHE4 Isoforms and Chip-seq annotation from Encode project are indicated.*

Sanger sequencing of the bisulfite PCR showed that CpG dinucleotide Chr3:5065165 is partially hypomethylated, as a result of *SMARCB1* re-expression (Figure 7.4). Moreover, of the five additional CpG dinucleotides identified, three were consistently hypomethylated in response to *SMARCB1* re-expression.

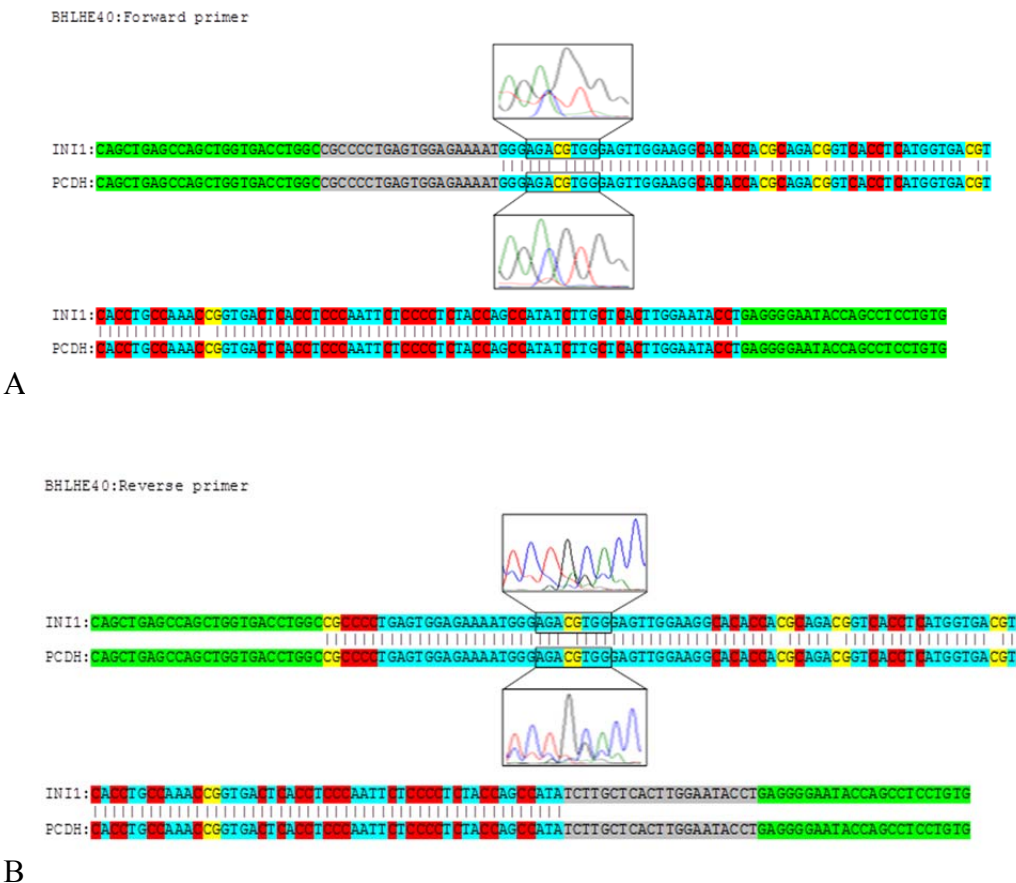
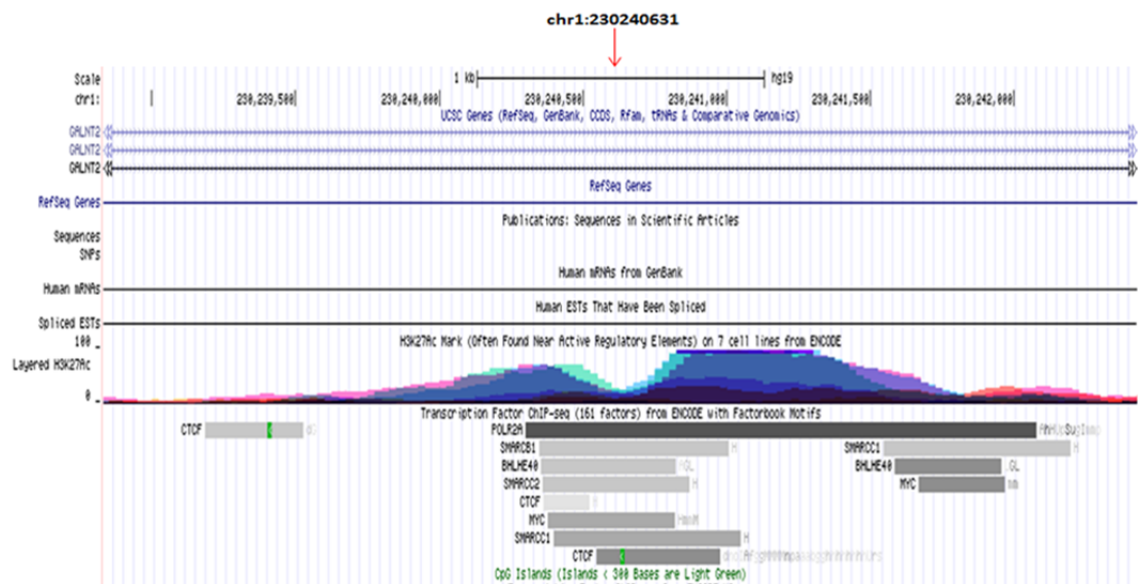


Figure 7.4 Sanger sequencing indicates demethylation of *BHLHE4* in response to *SMARCB1* re-expression in G401 cell line. Forward (A) and reverse primers (B) are indicated in green. Blue colour represents matches between *SMARCB1* and *PCDH*. Cytosine bases highlighted in red are unmethylated and converted into Ts and As in the forward and reverse strand, respectively. The trace is given for the CpG of interest (Chr3:5065165). Evidence for demethylation is indicated by overlapping traces.

*SMARCB1 re-expression induces changes in GALNT2 methylation status*

CpG dinucleotide of interest (Chr1:230240631) was analysed using the UCSC browser. The CpG of interest was located within the gene body upstream of the promoter (Figure 7.5). ChIP-Seq data from Encode indicates the binding of the core unit of the SWI/SNF complex to the region, as well as POLR2A and MYC. The location is characterised by a high levels of H3K27Ac, indicating that the CpG dinucleotide of interest is positioned within an active regulatory region.



*Figure 7.5 Transcription factor ChIP-Seq data for GALNT 2 derived from the Encode project. The CpG of interest is indicated by the red arrow. Different GALNT2 isoforms and Chip-seq annotation from Encode project are indicated.*

Sanger sequencing of the bisulfite PCR showed that CpG dinucleotide Chr1:230240631 is hypomethylated, as a result of *SMARCB1* re-expression (Figure 7.6). Also three other adjacent CpG dinucleotides all showed evidence of hypomethylation in each of the Malignant Rhabdoid Tumour cell lines in a *SMARCB1* dependent manner.

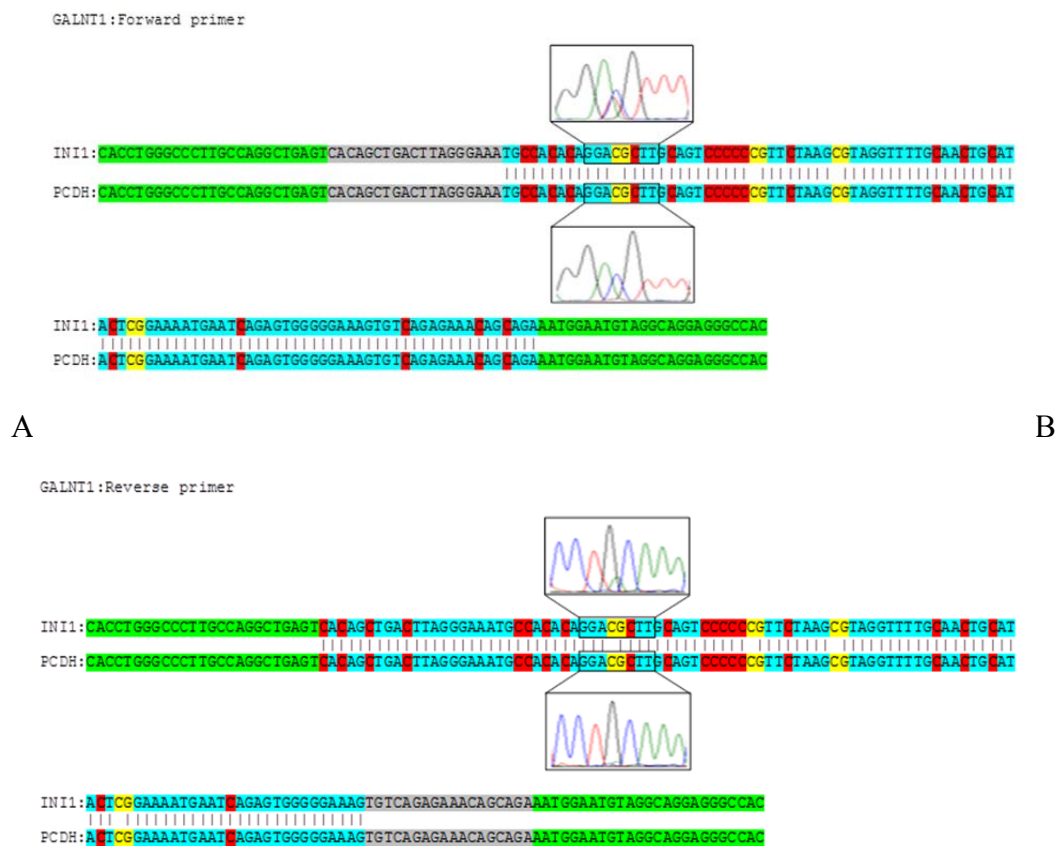


Figure 7.6 Sanger sequencing indicates demethylation of *GALNT2* in response to *SMARCB1* re-expression in G401 cell lines. Forward (A) and reverse primers (B) are indicated in green. Blue colour represents matches between *SMARCB1* and *PCDH*. Cytosine bases highlighted in red are unmethylated and converted into Ts and As in the forward and reverse strand, respectively. The trace is given for the CpG of interest (Chr3:5065165). Evidence for demethylation is indicated by overlapping traces.

### 7.3.2 Gene expression cross-referencing in model data identifies *SMARCB1* role in epigenetic mechanism

#### 7.3.2.1 Cross-referencing expression data in model data reveals demethylating effects of *SMARCB1* re-expression

Cross-referencing methylation array data between primary and model cell lines reveals that *SMARCB1* promotes demethylation of many CpG islands. In order to understand the mechanism of *SMARCB1* loss in the methylation-dependent modification of gene expression, expression data from *SMARCB1* re-expressing cells was compared with the expression data set obtained from Rhabdoid cells treated with 5-aza-2-deoxycytidine (5-azaCdR). In particular, we hypothesised that the re-expression of *SMARCB1* could be a direct effector of demethylation, in a similar manner to 5-aza-2-deoxycytidine.

Significantly differentially expressed genes from DEseq analysis from the two conditions were directly compared, revealing 14 genes which were deregulated after both *SMARCB1* re-expression and 5-azaCdR treatment (Figure 7.7 and Table 7.2).

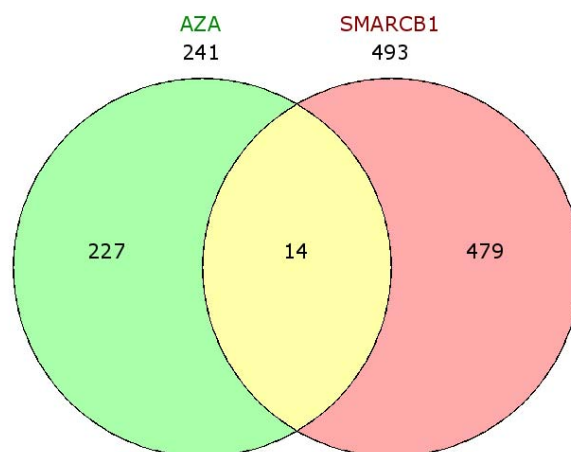


Figure 7.7 Venn diagram of differentially expressed genes as result of *SMARCB1* re-expression and 5-azaCdR treatment.

Gene Symbol	Adjusted <i>p</i> -value	log2fold change ( <i>SMARCB1</i> )
<b>DNA2</b>	1.21E-19	7.552249108
<b>PODXL</b>	7.46E-43	6.323432896
<b>SERPINE2</b>	1.85E-18	5.832168685
<b>PLK2</b>	7.77E-07	4.635228454
<b>ERC1</b>	4.23E-07	4.323596829
<b>PDLIM1</b>	2.61E-16	4.288680622
<b>C3</b>	1.08E-19	4.058441615
<b>HCLCS1</b>	4.10E-11	3.515754953
<b>DUSP5</b>	4.22E-07	2.508325668
<b>UBE4B</b>	0.000227	2.136383472
<b>ELL2</b>	1.08E-05	2.062217233
<b>ELF3</b>	3.58E-05	1.957602117
<b>B4GALNT3</b>	2.20E-05	1.852817531
<b>PDE4C</b>	9.32E-06	-1.705233083

*Table 7.2 List of the common differentially expressed genes resulting from SMARCB1 re-expression and 5-azaCdR data cross-referencing. Genes are ranked based on the moderated log2fold change in SMARCB1 re-expressing Rhabdoid cells.*

In this list two genes were also previously identified in this study as genes implicated in Rhabdoid tumorigenesis. *DNA2* was in the top 15 most differentially expressed following *SMARCB1* re-expression, while *PDE4C* is part of the family of PDE4 cyclic AMP (cAMP)-specific phosphodiesterase together with *PDE4D* (identified as significant deregulated genes in response to 5-azaCdR treatment). Interestingly, both *PDE4C* and *PDE4D* are two of the few genes showing downregulation following *SMARCB1* re-expression and 5-azaCdR treatment. Moreover, *SERPINE2* was indicated to be differentially expressed in primary tumours (Gadd, Sredni et al. 2010) and following induction of *SMARCB1* in Rhabdoid cell lines (Morozov, Lee et al. 2007).

### 7.3.2.2 Platform cross-referencing identifies genes directly demethylated by *SMARCB1*

We suspected that the small overlapping subsets of genes obtained in the previously comparative analysis was a result of the broad and global effect on methylation of 5-azaCdR treatment but a somewhat more modest effect on expression than *SMARCB1* re-expression. To better understand the correlation between gene expression and changes in methylation, cross-referencing analysis between RNA-seq and 450K data in Rhabdoid cell re-expressing *SMARCB1* was attempted. Significant differentially expressed genes ( $p$ -value  $<0.05$ ; moderate log Fold Change  $<1$  and  $>1$ ; raw count per gene with  $\geq 50$  counts in at least 3 cell lines) were cross-referenced with significantly differentially methylated CpG probes (delta values  $>0.3$  and  $<-0.3$  in at least 3 cell lines).

From the genes which satisfied these criteria four were selected for further investigation. *GLI2*, *GLI3*, *JARID2* and *NTN4* genes were selected by a literature search for those genes of particular relevance to cancer, and supported by transcription factor ChIP-Seq data from Encode, using UCSC Genome Browser. *GLI2* regulates the Sonic Hedgehog (Shh) signalling pathway and promotes metastasis in prostate and melanoma cell lines; moreover *GLI1* is reported down-regulated following re-expression of *SMARCB1* (Thiyagarajan, Bhatia et al. 2007, Alexaki, Javelaud et al. 2010, Jagani, Mora-Blanco et al. 2010). Overexpression of *GLI3*, another member of the Shh pathway, is associated with tumourigenesis in colon cancer and with poor survival in non-small cell lung cancer patients (Bai, Lin et al. 2013, Iwasaki, Nakano et al. 2013). *JARID2* is involved in transcriptional repression in leukaemia and promotes maintenance of undifferentiated myogenic phenotype in rhabdomyosarcomas (Puda, Milosevic et al. 2012, Walters, Villarejo-Balcells et al. 2014). *NTN4* is a *SMARCB1*-related molecule which inhibits angiogenesis and tumour cell growth in colorectal carcinoma, also *NTN4* downregulation is observed in glioblastoma promoting cell proliferation (Hu, Ylivinkka et al. 2012, Eveno, Contreres et al. 2013).



The mRNA level of these genes was analysed by quantitative RT-PCR following *SMARCB1* re-expression and compared with mRNA of Rhabdoid cells treated with 5-azaCdR. We observed a modest increase of *GLI2* expression upon *SMARCB1* re-expression (Figure 7.8) in A204 and CHLA-266 (14 and 33-fold, respectively), while a more marked upregulation was noticeable in G401 and STA-WT-1 (5545 and 141-fold, respectively); similar upregulation was observed after 5-azaCdR treatment in G401, A204 and STA-WT-1 (14.8 to 45.8-fold), whereas induction was more modest in CHLA (4.2-fold). Expression of *GLI3* following *SMARCB1* induction (Figure 7.8) was substantial in G401 (636.5-fold), whilst more modest increases were observed in other cell lines (6.8 to 15.8-fold) Figure 7.9; a small upregulation of *GLI3* upregulation was observed upon 5-azaCdR treatment (1.2 to 4.4-fold). *JARID2* expression (Figure 7.9) was modestly increased in A204, STA-WT-1 and CHLA-266 (1.5 to 7.4-fold) in comparison to G401 (77.5-fold); a subtle upregulation in all cells were observed in response to 5-azaCdR treatment (1.2 to 1.7-fold). In response to *SMARCB1* re-expression *NTN4* (Figure 7.9) expression was increased in all cells (2.3 to 17.7-fold) with significant over-expression in G401 (1694.9-fold); a very modest upregulation upon 5-azaCdR treatment was observed (1.5 to 2.4-fold), except for G401 where downregulation was detected (0.8-fold).

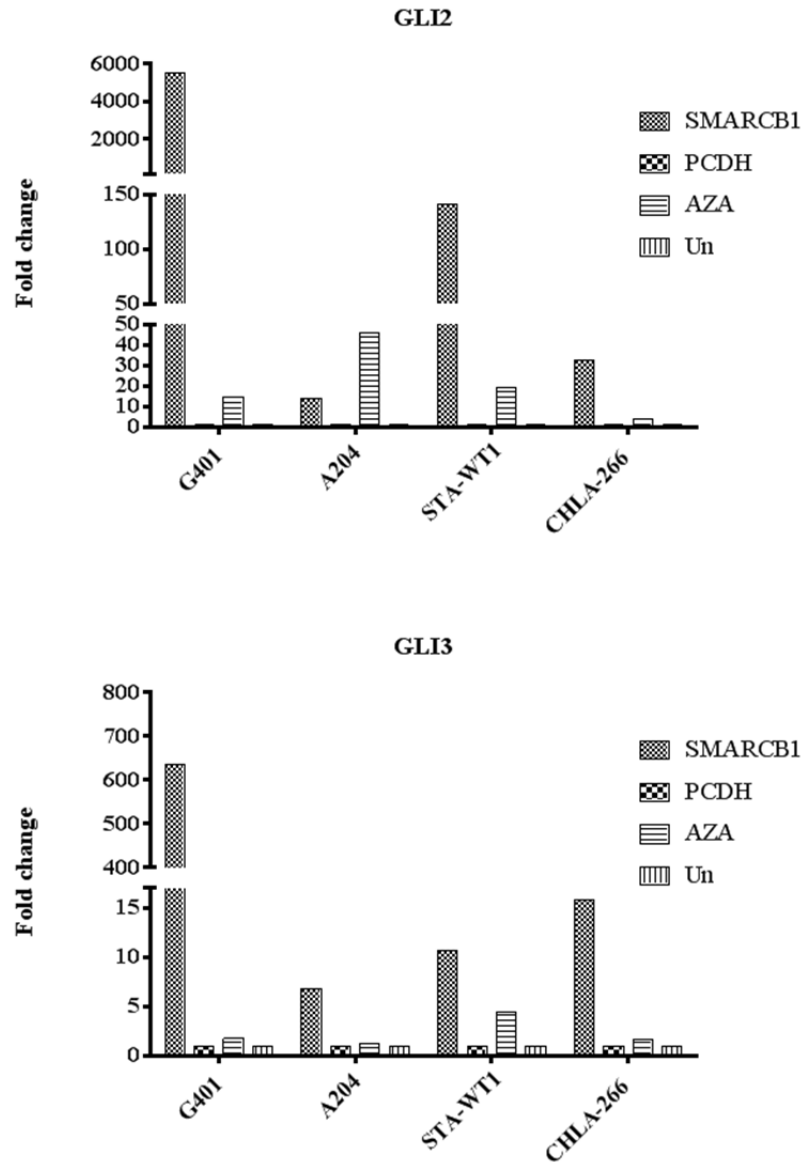


Figure 7.8 SMARCB1 re-expression induces upregulation of GLI2 and GLI3 as a result of demethylation and/or SMARCB1 re-expression. The mRNA levels were measured by real-time RT-PCR and normalised to beta-2-microglobulin. Values are given relative to untreated cell lines (Un) or empty vector controls (pCDH) as appropriate and were the mean of 3 independent measurements. Rhabdoid cell re-expressing SMARCB1 are indicated with SMARCB1, while cell treated with 5-azaCdR with AZA.

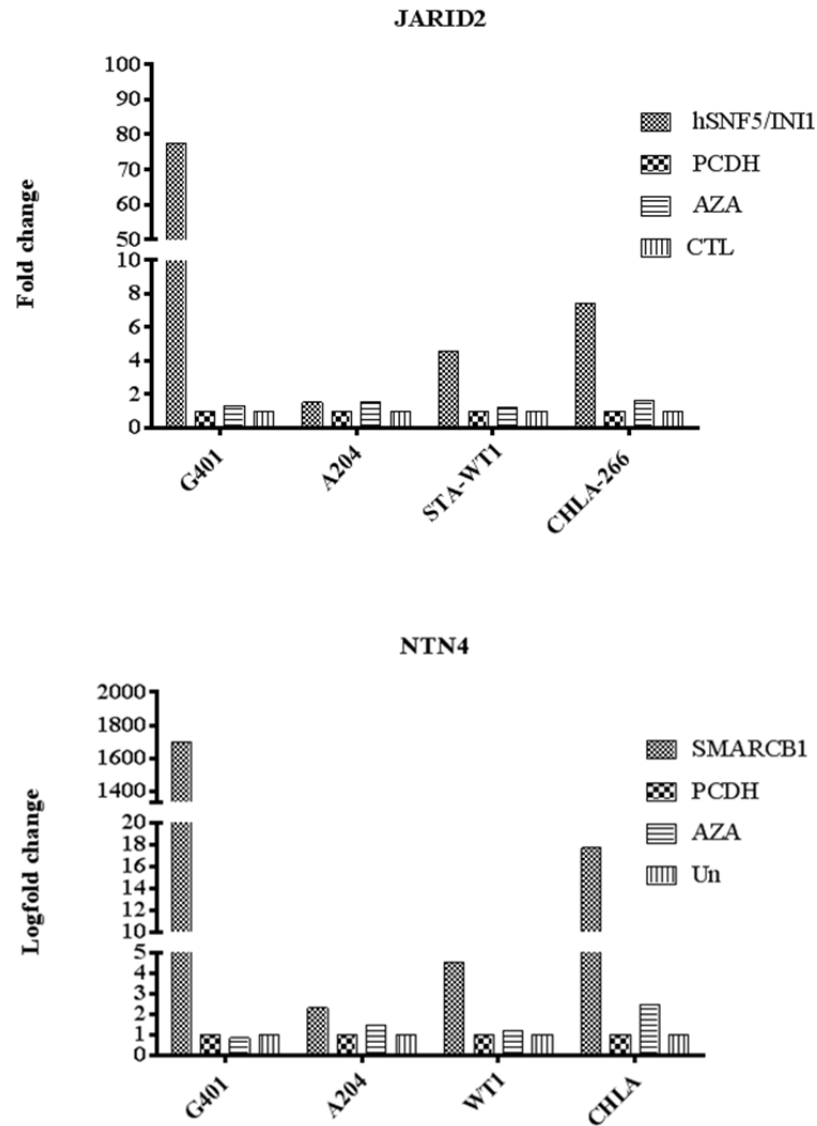


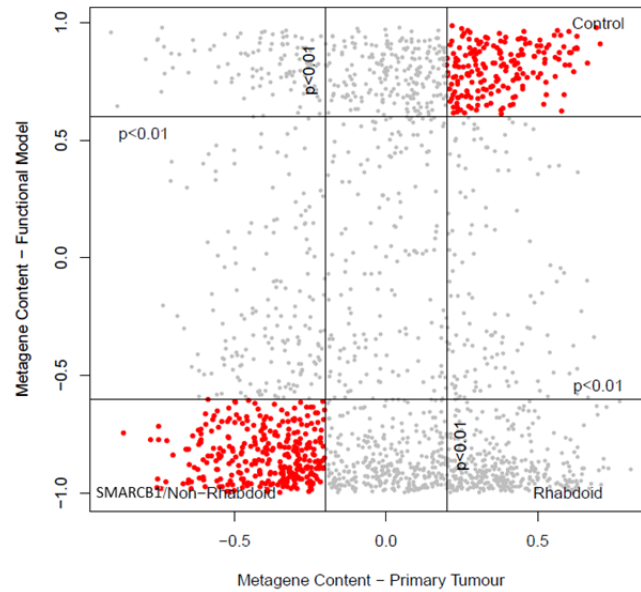
Figure 7.9 SMARCB1 re-expression induces upregulation of JARID2 and NTN4, as a result of demethylation and/or SMARCB1 re-expression. The mRNA levels were measured by real-time RT-PCR and normalised to beta-2-microglobulin. Values are given relative to untreated cell lines (Un) or empty vector controls (pCDH) as appropriate and were the mean of 3 independent measurements. Rhabdoid cell re-expressing SMARCB1 are indicated with SMARCB1, while cell treated with 5-azaCdR with AZA.

## 7.4 Pathway analysis

### 7.4.1 Cross-referencing of multiple gene expression gene sets identifies gene signatures directly controlled by *SMARCB1*

In order to understand in an analytical and systematic manner the gene signature(s) involved in Rhabdoid tumorigenesis, we cross-referenced the expression data set derived from three analyses (Primary Tumours, Rhabdoid Tumours vs Medulloblastoma; Cell line models, *SMARCB1* re-expression vs empty vector controls; 5-aza-CdR treatment, Treated vs untreated) specifically the particular genes which constituted and defined the metagenes calculated for each experimental condition. The contribution of each gene to a particular defining metagene signature was calculated by means of a Pearson correlation coefficient of the expression of each gene with the values for each metagene.

First, the metagene expression signatures discriminating Rhabdoid primary tumour and Medulloblastoma were extracted from the expression data and the metagene content with respect to individual genes was calculated. This gave a score for each gene of -1 to 1 where 1 indicates perfect positive correlation of a particular gene with metagene and -1 perfect anti-correlation. An adjusted p-value was calculated for the correlation of each gene with the metagene. The same process was repeated with a metagene signature defining the difference between *SMARCB1* re-expressing Rhabdoid Tumour cell lines. The analysis was focused on finding shared genes between *SMARCB1* re-expressing cells and Medulloblastoma (which in this analysis corresponds to non Rhabdoid phenotype tumours) and vice versa between vector control infectants and Rhabdoid primary tumours (which represent the Rhabdoid phenotype) (Figure 7.10). Figure 7.10 shows in red the genes which form a significant ( $p < 0.01$ ) part of both the primary and cell line model. This analysis is far more sensitive than simple cross-referencing of two separate DESeq analyses.



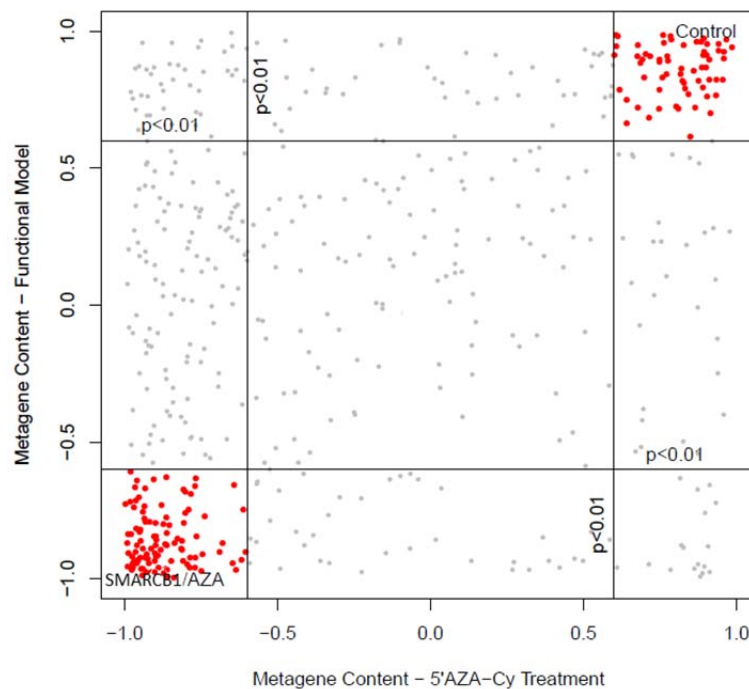
*Figure 7.10 Cross-referencing of two metagene signatures from Malignant Rhabdoid Tumours Primary tumours (Rhabdoid) vs Medulloblastomas (Non-Rhabdoid) and cell model data re-expressing SMARCB1 and control infectants reveals common significantly deregulated genes  $p < 0.01$  (red). On the x-axis are plotted the values for metagene content (correlation of individual genes with metagene values) extracted from primary data, while the y-axis gives metagene content from the cell line models.*

Data integration of metagene content identified 187 common genes that defined a Rhabdoid/*SMARCB1* negative phenotype across the data sets. Table 7.3 lists the top 20 significant genes as a result of this cross-referencing and ranked by metagene content in *SMARCB1* re-expressing cells. Complete list in Appendix I.

Gene Symbol	Metagene content Medulloblastomas	Metagene content SMARCB1 re-expressing cells
<b>IDH1</b>	0.218190152	0.985473815
<b>SUCLG2</b>	0.69429913	0.97669975
<b>TYSND1</b>	0.249656719	0.97376244
<b>RARG</b>	0.565913736	0.966783287
<b>RGS1</b>	0.269581376	0.964187488
<b>HIC1</b>	0.603438878	0.963633075
<b>CREG1</b>	0.454614525	0.962206288
<b>HMHA1</b>	0.276917082	0.959244273
<b>LHX6</b>	0.257225532	0.954712178
<b>TGIF2</b>	0.373928224	0.952033316
<b>TNFAIP8L1</b>	0.555467972	0.947460948
<b>GNPNAT1</b>	0.575562612	0.945400416
<b>PIM3</b>	0.251816526	0.940025544
<b>ELFN1</b>	0.328714614	0.937885973
<b>ALDH16A1</b>	0.395197109	0.937582613
<b>GAREML</b>	0.334710923	0.934206727
<b>LINC00982</b>	0.445616696	0.933444824
<b>CLN3</b>	0.273734686	0.931034387
<b>SIX5</b>	0.458252972	0.929795247

*Table 7.3 Top 20 genes that define the Non- Rhabdoid phenotype resulting from cross-referencing Medulloblastomas metagene content and SMARCB1 re-expressing cell lines metagene content. Each gene constitutes a significant part of the defining metagene signature of both primary and functional cell line models.*

The same analysis was performed by comparing the metagene signature of Rhabdoid cells re-expressing *SMARCB1* with the metagene content of Rhabdoid cells treated with 5-azaCdR. The analysis was focused on finding correlations that define a *SMARCB1*/methylation-dependent from both data sets (Figure 7.11). Data integration identified 64 common genes strongly correlated with metagenes corresponding to *SMARCB1* re-expressing cells and 5-azaCdR treated cells. Table 7.3 lists the top 20 significant genes as a result of this cross-referencing and ranked by metagene content in *SMARCB1* re-expressing cells. Complete list in Appendix I.



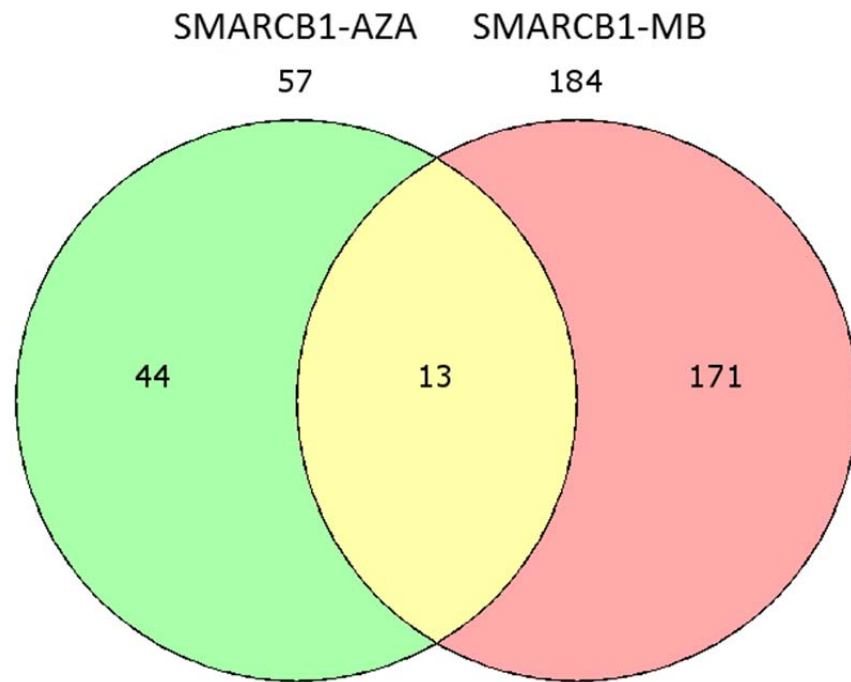
*Figure 7.11 Cross-referencing of metagene signatures from Rhabdoid cells re-expressing SMARCB1 (SMARCB1) and from Rhabdoid cells treated with 5-azaCdR (AZA) reveals a common gene signature in the data set (in red). On the x-axis are plotted the values for metagene content of each gene extracted from cells treated with 5-azaCdR, the y-axis shows the metagene content values from the SMARCB1 model data set.*

<b>Gene Symbol</b>	<b>Metagene content 5-azaCdR treated cells</b>	<b>Metagene content SMARCB1 re-expressing cells</b>
<b>IDH1</b>	0.940223791	0.985473815
<b>FGFBP3</b>	0.967750489	0.96766879
<b>FAM19A5</b>	0.900312294	0.958655837
<b>BMI1</b>	0.922384221	0.958286584
<b>SPAG6</b>	0.822514539	0.934151603
<b>ZSCAN10</b>	0.701066641	0.914554305
<b>H2AFY2</b>	0.82244248	0.911040681
<b>RAD51B</b>	0.952106848	0.903970282
<b>IFI35</b>	0.762395879	0.903854604
<b>DLGAP1-AS3</b>	0.911638216	0.902348866
<b>DBP</b>	0.867467803	0.900689488
<b>DNM3OS</b>	0.925276636	0.894459127
<b>RPRM</b>	0.971562868	0.893505765
<b>CA11</b>	0.91068111	0.892449763
<b>CPNE5</b>	0.794414076	0.890840559
<b>SH2B2</b>	0.867731083	0.886856403
<b>KLF15</b>	0.959252724	0.884368174
<b>CCDC74B</b>	0.819681347	0.875987288
<b>HSF4</b>	0.721212585	0.86557006

*Table 7.4 Top 20 genes that defined the Non-Rhabdoid phenotype resulting from cross-referencing 5-azaCdR treated cells metagene content and SMARCB1 re-expressing cell lines metagene content. Metagene content score is indicated.*



Finally the two gene lists obtained from this analysis was compared again: 13 genes common in both lists were identified (Figure 7.12 and Table 7.5).



*Figure 7.12 Venn diagram of Methylation-dependent Rhabdoid gene signatures common between SMARCB1 re-expressing cells and Medulloblastomas (SMARCB1-MB) and between SMARCB1 re-expressing cells and 5-azaCdR treated cells.*

Gene Symbol	Metagene content 5-azaCdR treated cells	Metagene content SMARCB1 re-expressing cells	Metagene content Medulloblastomas
<b>CALY</b>	0.892541	0.776439	0.306465
<b>CCDC74B</b>	0.819681	0.875987	0.285176
<b>DNM3OS</b>	0.925277	0.894459	0.628264
<b>IDH1</b>	0.940224	0.985474	0.21819
<b>IFI35</b>	0.762396	0.903855	0.537416
<b>KAZALD1</b>	0.950182	0.827627	0.664685
<b>KIAA0101</b>	0.747337	0.641398	0.221889
<b>MACROD1</b>	0.969651	0.787554	0.204009
<b>NQO2</b>	0.831308	0.757666	0.369473
<b>PBK</b>	0.909212	0.676959	0.303351
<b>PHGDH</b>	0.614676	0.848927	0.379996
<b>PRDM16</b>	0.955471	0.76299	0.429157
<b>SPC24</b>	0.807937	0.828866	0.35925

*Table 7.5 Non Rhabdoid gene signatures common between SMARCB1 re-expressing cells and Medulloblastomas (SMARCB1-MB) and between SMARCB1 re-expressing cell sand 5-azaCdR treated cells. Metagene content score is indicated.*

*CCDC74B* alters regulation of mitosis during initiation as well as S-phase and G2-M progression (Yao, Bee et al. 2010). *Dnm3os* is a gene transcribed as a non-coding RNA (ncRNA) that contains three micro RNAs (miRNAs), miR-199a, miR-199a, and miR-214; it is indispensable for normal skeletal development and body growth in mammals (Watanabe, Sato et al. 2008). Mutation of *IDH1* is frequent in gliomas, although the oncogenic mechanism is still not understood (Cohen, Holmen et al. 2013). *KAZALD1* is highly methylated in sarcomatoid-type peritoneal mesothelioma (Hama, Watanabe et al. 2012). *KIAA0101* protein is overexpressed in breast cancer and modifies directly centrosomes (Kais, Barsky et al. 2011). *BK/TOPK* (PDZ-binding kinase, T-LAK-cell-originated protein kinase) is a serine-threonine kinase that is overexpressed in a variety of tumour cells but whose role in oncogenesis remains unclear (Hu, Gartenhaus et al.

2010). *PRDM16* gene is known to be rearranged in Acute Myeloid Leukaemia (AML) and Myelodysplasia Syndrome (MDS) (Duhoux, Ameye et al. 2012). Ndc80/Hec1 assembles as part of a 4-subunit Ndc80 complex that includes Nuf2, Spc24, and Spc25. Ndc80/Hec1 is often upregulated in tumours and mice overexpressing Ndc80 show increase in cancer formation (Diaz-Rodriguez, Sotillo et al. 2008).

#### **7.4.2 Gene set enrichment analysis**

Gene set enrichment analysis (GSEA) is a computational method that identifies statistically significant differences in gene expression between two conditions of an *a priori* defined list of genes (gene set). GSEA requires a ranking of genes defined according to two conditions. The GSEA software compares the position of genes in a gene set within the predefined gene ranking such that a non-random distribution within the ranking would indicate significant enrichment of that gene set. Gene sets are typically taken from a molecular signature database i.e. MSigDB based on published data or recognised gene ontologies.

The previous unsupervised RNA-seq metagene analyses were taken as the basis of the pre-defined gene ranking. Genes were ranked based on the degree of association (Pearson correlation coefficient) to the representative metagene. GSEA was performed on these pre-ranked gene lists for all gene expression analysis: primary tumours and functional cell models. The parameters were set to identify significantly enriched gene sets with a minimum of 15 genes and a maximum of 1000 genes. The GSEA tool computed the level of significance (p-value) for each enriched gene set, and a false discovery rate (FDR). The enriched gene sets were ranked according to the p-value and FDR, and gene sets with a p-value less than 0.05 and an FDR rate less than 0.25 were excluded from the results. Using such cut-off values resulted in the selection of only those gene sets with highly significant gene enrichment.

#### *7.4.2.1 GSEA analysis of differentially regulated signalling pathways in primary Rhabdoid Tumours*

GSEA analysis was performed to identify biological processes or cancer signalling pathways that are enriched specifically in Rhabdoid Tumours. The analysis was performed upon a pre-ranked list of genes ranked according to their correlation with the primary Rhabdoid tumour metagene extracted from RNA-seq data. This ranked gene list was analysed against the curated gene set database MSigDB C2 (an annotated Gene set collected from various sources such as online pathway databases, publications in PubMed, and knowledge of domain experts.) and MSigDB C6 (an oncogenic signature annotation) (<http://www.broadinstitute.org/gsea/index.jsp>).

GSEA analysis against the C2 gene set library reported 1117 out of 3255 gene sets presenting an FDR < 25% and  $p$ -value <1% and enriched when comparing primary Rhabdoid Tumours with Medulloblastoma. Also 26 gene sets were also significantly de-enriched. Selected upregulated gene sets (positively enriched for Rhabdoid Tumours ) are indicated in Table 7.6 and Figure 7.13. These include genes commonly down-regulated in human alveolar rhabdomyosarcomas, genes co-ordinately up-regulated in mouse embryonic stem cells (ESC) and genes involved in p53-Independent G1/S DNA damage checkpoint. Selected gene sets significantly de-enriched i.e. were downregulated or negatively enriched for Rhabdoid Tumours are indicated in Table 7.6 and Figure 7.14. These include genes down-regulated and associated with poor survival of patients with metastatic Neuroblastoma lacking *MYCN* amplification; and genes up-regulated during neuronal differentiation of SH-SY5 neuroblastoma cells in response to stimulation by tretinoin.

Gene set	ES	FDR	p-val
<b>REN_ALVEOLAR_ RHABDOMYOSARCOMA_DN</b>	0.586	0	0
<b>WONG_EMBRYONIC_ STEM_CELL_CORE</b>	0.563	0	0
<b>REACTOME_P53_INDEPENDENT_G1_ S_DNA_DAMAGE_CHECKPOINT</b>	0.772	0	0
<b>ASGHARZADEH_NEUROBLASTOMA_ POOR_SURVIVAL_DN</b>	-0.635	6.41E-05	0.001
<b>LE_NEURONAL_DIFFERENTIATION_ UP</b>	-0.655	0	0.0019
<b>KIM_ALL_DISORDERS_OLIGODENDR OCYTE_NUMBER_CORR_DN</b>	-0.536	0.0039	0.030

*Table 7.6 Selected C2 GSEA pathways upregulated and downregulated in primary Malignant Rhabdoid Tumours. Relative enrichment score (ES), false discovery rate (FDR) and p-values are also displayed. For details of gene set contents see MSigDB.*

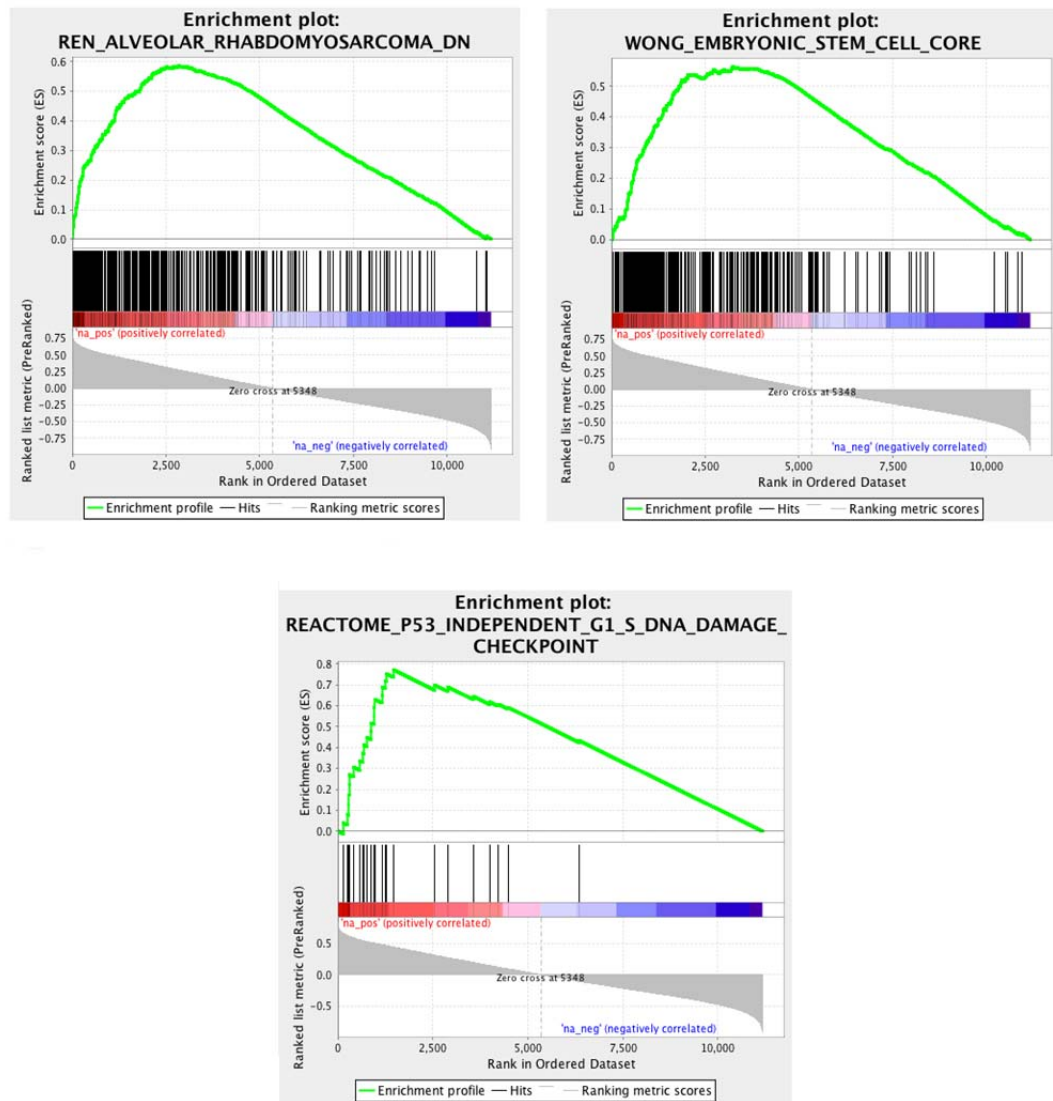


Figure 7.13 GSEA enrichment plots of gene sets: Ren alveolar rhabdomyosarcomas, Wong embryonic stem cells core and Reactome p53 independent G1/S DNA damage checkpoint pathways genes in Malignant Rhabdoid Tumour (red) vs Medulloblastomas (blue). Genes in all three pathways showed significant enrichment in Malignant Rhabdoid Tumours. The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas value of the ranking metric for each gene in the gene set is given by a vertical black tick.

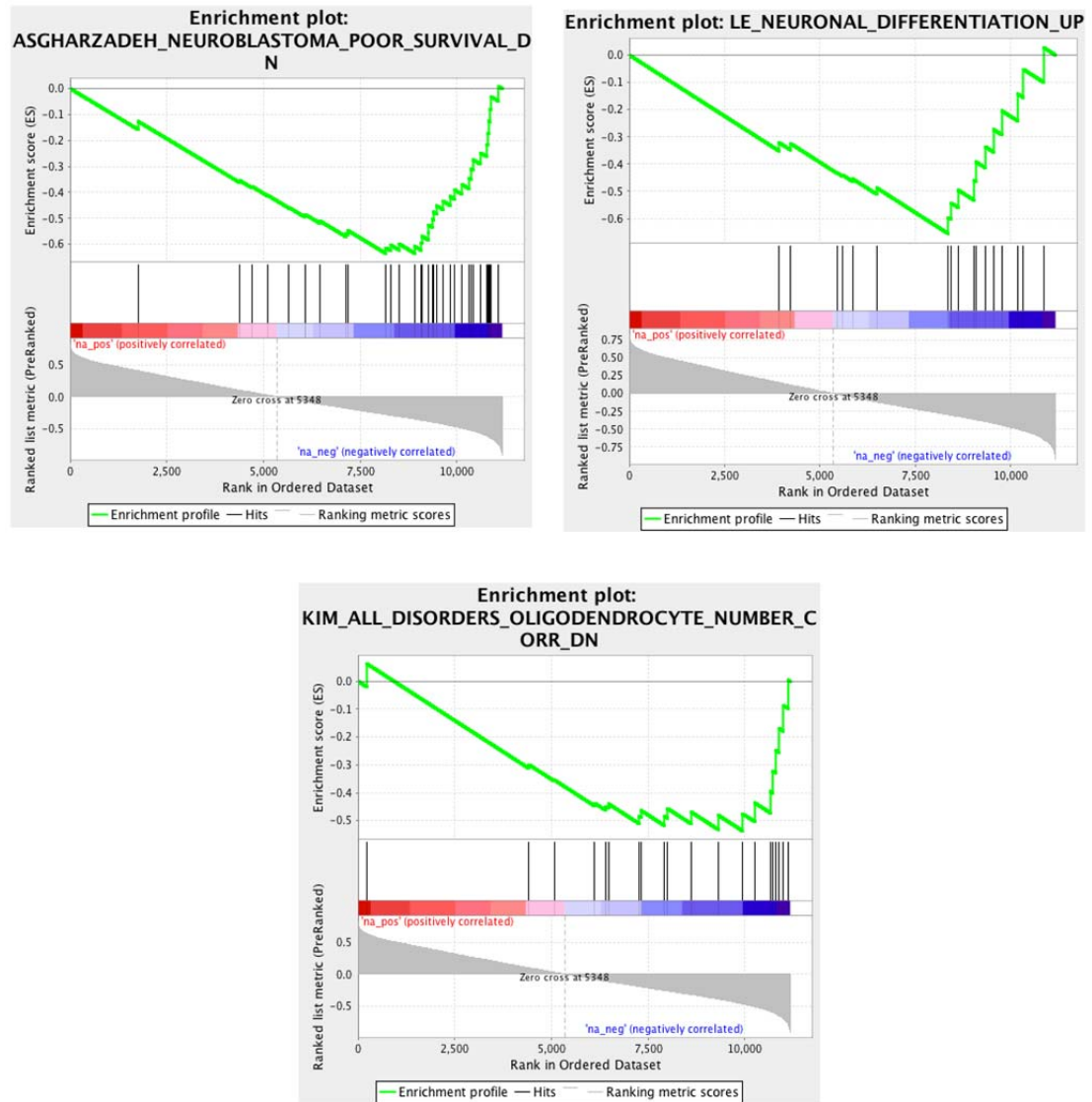


Figure 7.14 GSEA enrichment plot of Asgharzadeh neuroblastoma poor survival, Le neuronal differentiation, Kim all disorders oligodendrocyte number corr DN in Malignant Rhabdoid Tumour (red) vs Medulloblastomas (blue). Genes in all three pathways showed significant negative enrichment in Malignant Rhabdoid Tumours. The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas value of the ranking metric for each gene in the gene set is given by a vertical black tick.

Analysis against the C6 MSigDB library identified 37 out of 187 genes set significantly upregulated and 26 significantly downregulated (FDR < 25% and *p*value <1%) when comparing primary Rhabdoid Tumours with Medulloblastoma.

Selected upregulated gene sets (gene positively enriched for Rhabdoid Tumours) are indicated in (Table 7.7 and Figure 7.15): these include genes up-regulated in kidney epithelium cells over-expressing *GLI1*; genes up-regulated in breast cancer cells over-expressing *YAP*. Gene sets negatively enriched include genes up-regulated in neurons (see Table 7.7 and Figure 7.15);.

Gene set	ES	FDR	<i>p</i> -val
<b>GLI1_UP.V1_UP</b>	0.412	0.046	0.048
<b>YAP1_UP</b>	0.403	0.020	0.014
<b>CAHOY_NEURONAL</b>	-0.511	0	0

*Table 7.7 3 selected C6 GSEA pathways significantly upregulated and downregulated in Rhabdoid Tumours when compared with Medulloblastomas. Relative enrichment score (ES), false discovery rate (FDR) and p-value are also displayed.*



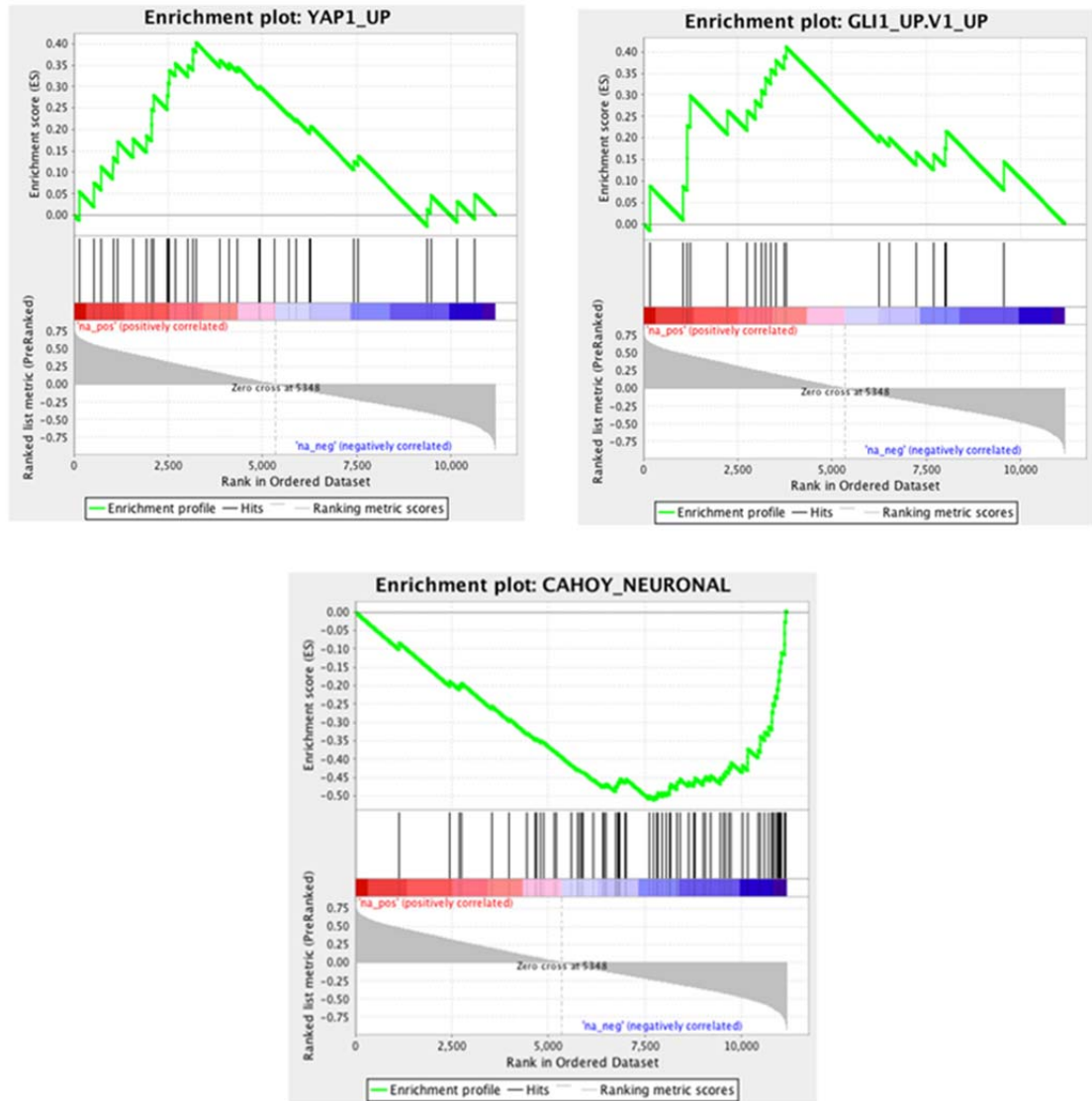


Figure 7.15 GSEA enrichment plot of YAP1, GLI1 upregulated pathways genes and CAHOY NEURONAL downregulated pathways in Malignant Rhabdoid Tumour (red) vs Medulloblastomas (blue). The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas value of the ranking metric for each gene in the gene set is given by a vertical black tick.

GSEA identified pathways in accordance with that previously described in the comparative analysis and to the literature. Analysis based on the C2 library underlines the substantial phenotypic difference between Malignant Rhabdoid and Medulloblastoma tumours. Significantly deregulated genes in Rhabdoid Tumours present positive enrichment for genes commonly down-regulated in human Rhabdomyosarcomas, up-regulated in mouse embryonic stem cells and involved in p53-Independent G1/S DNA damage checkpoint. This analysis is in accord with previous reports that underline the similarity between the Malignant Rhabdoid Tumours and Rhabdomyosarcomas (Armeanu-Ebinger, Herrmann et al. 2012). In fact both malignancies show distinctive expression of embryonic stem cell markers (Deisch, Raisanen et al. 2011) and the upregulation of cell cycle check point genes such as *NOXA* in a *p53* independent manner (Kuwahara, Wei et al. 2013). Moreover, previous comparative analysis between Medulloblastomas and Rhabdoid Tumours showed a substantive difference in neuronal differentiation phenotype (Higashino, Narita et al. 2003).

Interesting both *GLII* and *YAP1* gene sets are observed to be significantly enriched in the c6 library. This again reflects previous observation in Rhabdoid cell lines and in Atypical Teratoid Rhabdoid primary Tumours, where *GLII* and *YAP1* are upregulated respectively (Jagani, Mora-Blanco et al. 2010, Jeibmann, Eikmeier et al. 2014).

#### 7.4.2.2 GSEA analysis distinguishes differentially regulated signalling pathways in Extra Cranial Rhabdoid Tumours and Atypical Teratoid Rhabdoid Tumours

A GSEA analysis was performed by analysing a list of genes ranked according to their correlation with a metagene which defined the Extra Cranial Rhabdoid Tumour (ECRT) as compared to the Atypical Teratoid Rhabdoid Tumours (ATRT). This pre-ranked list of genes was analysed against the C2 and C6 MSigDB gene set libraries.

Analysis against the C2 library identified 596 significantly enriched genesets (FDR < 25% and *p*-value <1%) out of 3201 and 257 significantly negatively enriched when comparing Atypical Teratoid Rhabdoid Tumours to in Extra Cranial Rhabdoid Tumours.

Gene set	ES	FDR	<i>p</i> -val
WONG_EMBRYONIC_STEM_CELL_CORE	0.5830646	0	0
LEE_TARGETS_OF_PTCH1_AND_SUFU_DN	-0.5977889	0	0

*Table 7.8 Two C2 library GSEA pathways upregulated and downregulated in Atypical Teratoid Rhabdoid Tumours when compared with Extra-cranial Rhabdoid Tumour. Relative enrichment score (ES), false discovery rate (FDR) and *p*-value are also displayed.*

Interestingly, the Embryonic stem cell core pathway previously indicated to be significantly enriched in the GSEA analysis of primary Rhabdoid Tumours against Medulloblastoma was more positively enriched for Atypical Teratoid Rhabdoid Tumours and therefore more negatively for Extra Cranial Rhabdoid Tumours. Moreover, Extra Cranial Rhabdoid Tumours showed enrichment of a gene set describing downregulation of targets of *PATCH1* and *SUFU* (Table 7.8 and Figure 7.16).

These results confirm previous observations: Extra Cranial Rhabdoid Tumours. and Atypical Teratoid Rhabdoid Tumours express many common stem cell-associated transcription factors, however the expression in Atypical Teratoid Rhabdoid Tumours is higher than the Extra Cranial malignancies (Deisch, Raisanen et al. 2011, Venneti, Le et al. 2011); *PTCH1* collaborates with *GLII* in modulation of the SHH pathway and is downregulated in A204 and G401 Rhabdoid cell lines by inhibition of EZH2 (Knutson, Warholic et al. 2013).

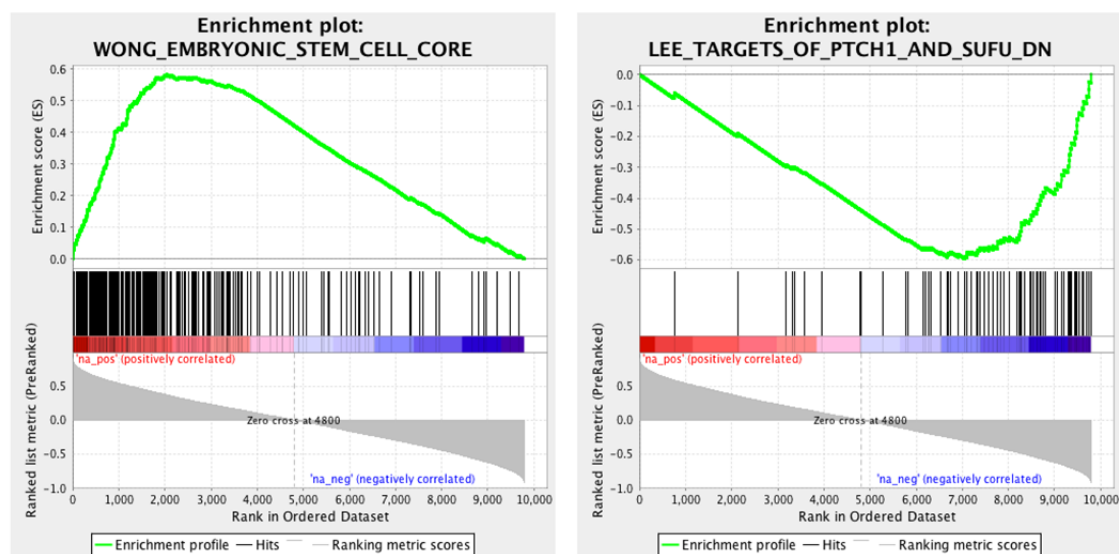


Figure 7.16 GSEA enrichment plot of two genes sets Wong Embryonic stem cell core gene (significantly enriched) and the Lee *PTCH1* and *SUFU* (significantly negatively enriched) in Atypical Teratoid Rhabdoid Tumours (red) vs Extra Cranial Rhabdoid Tumour (blue). The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas value of the ranking metric for each gene in the gene set is given by a vertical black tick.

Analysis of the Atypical Teratoid Rhabdoid Tumours vs Extra Cranial Rhabdoid Tumours metagene against the C6 MSigDb library identified 20 significantly enriched gene sets (FDR < 25% and  $p$ value <1%) out of 186 gene sets and 35 significantly negatively enriched. These included significant negative enrichment of a set of genes associated with the upregulation of KRAS in kidney i.e. positively enriched for Extra Cranial Rhabdoid Tumours. The Cahoy neuronal gene set was significantly negatively enriched in Extra Cranial Rhabdoid Tumours (Table 7.9 and Figure 7.17).

Gene set	ES	FDR	$p$ -val
<b>KRAS.KIDNEY_UP.V1_UP</b>	-0.680105	0	0
<b>CAHOY_NEURONAL</b>	-0.560470	0	0

*Table 7.9 C6 GSEA pathways upregulated and downregulated in Atypical Teratoid Rhabdoid Tumours when compared with Extra-cranial Rhabdoid Tumour. Relative enrichment score (ES), false discovery rate (FDR) and  $p$ -value are also displayed.*

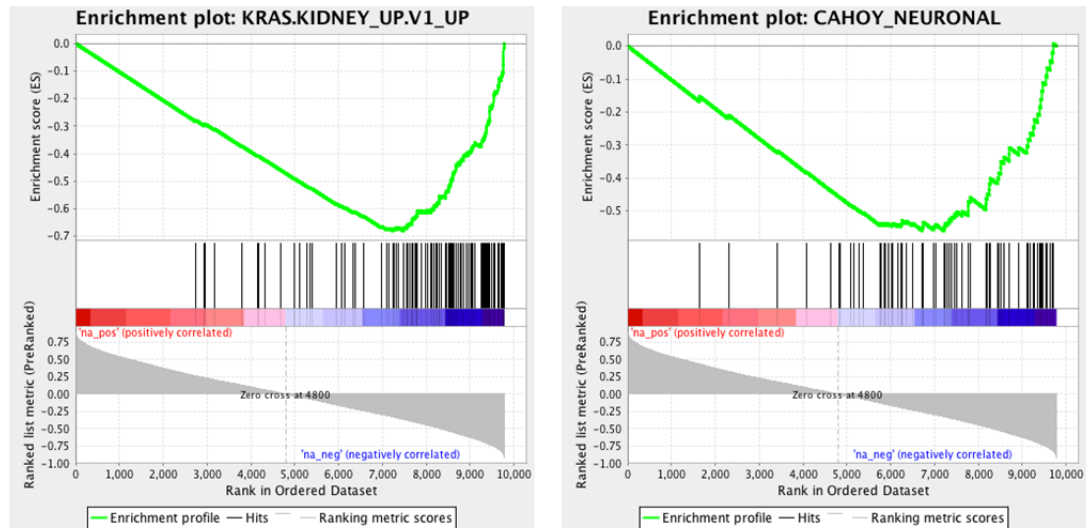


Figure 7.17 GSEA enrichment plot of KRAS pathways and Cahoy neuronal pathway in Atypical Teratoid Rhabdoid Tumours (red) vs Extra Cranial Rhabdoid Tumour (blue). The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas value of the ranking metric for each gene in the gene set is given by a vertical black tick.

#### 7.4.2.3 GSEA analysis of differentially regulated signalling pathways in Rhabdoid cell lines following *SMARCB1* re-expression

GSEA analysis was performed by analysing a pre-ranked list of genes correlated with a metagene that defines the difference between empty vector controls and *SMARCB1* re-expressing Rhabdoid cell lines. This pre-ranked list was analysed by GSEA against MSigDB libraries C2 and C6.

Analysis against the C2 library identified 221 significantly enriched gene sets FDR < 25% and *p*-value <1% out of 596 and 323 significantly negatively enriched following *SMARCB1* re-expression. The analysis indicated a wide number of up-regulated pathways involved in cell cycle and differentiation and also a significant enrichment of genes with promoters bound by *E2F4* (Figure 7.18 and Table 7.10). Significant negative enrichment was observed for the alveolar Rhabdomyosarcoma gene set pathway, *TGFB1* target pathway and the PDEF target genesets after knockdown of *PDEF* (Table 7.10 and Figure 7.19 ).

Gene set	ES	FDR	<i>p</i> -val
<b>MARSON_BOUND_BY_E2F4_UNSTIMULATED</b>	0.6865080	0	0
<b>GOBERT_OLIGODENDROCYTE_DIFFERENTIATION_UP</b>	0.5937639	0	0
<b>REACTOME_CELL_CYCLE</b>	0.6538154	0	0
<b>REN_ALVEOLAR_RHABDOMYOSARCOMA_DN</b>	-0.4799	0	0
<b>MCBRYAN_PUBERTAL_TGFB1_TARGETS_UP</b>	-0.50800	0	0
<b>GU_PDEF_TARGETS_UP</b>	-0.53733	2.51E-4	0

Table 7.10 Selected C2 GSEA pathways significantly enriched and negatively enriched in empty vector controls compared to *SMARCB1* re-expressing cell lines. Relative enrichment score (ES), false discovery rate (FDR) and *p*-value are also displayed.

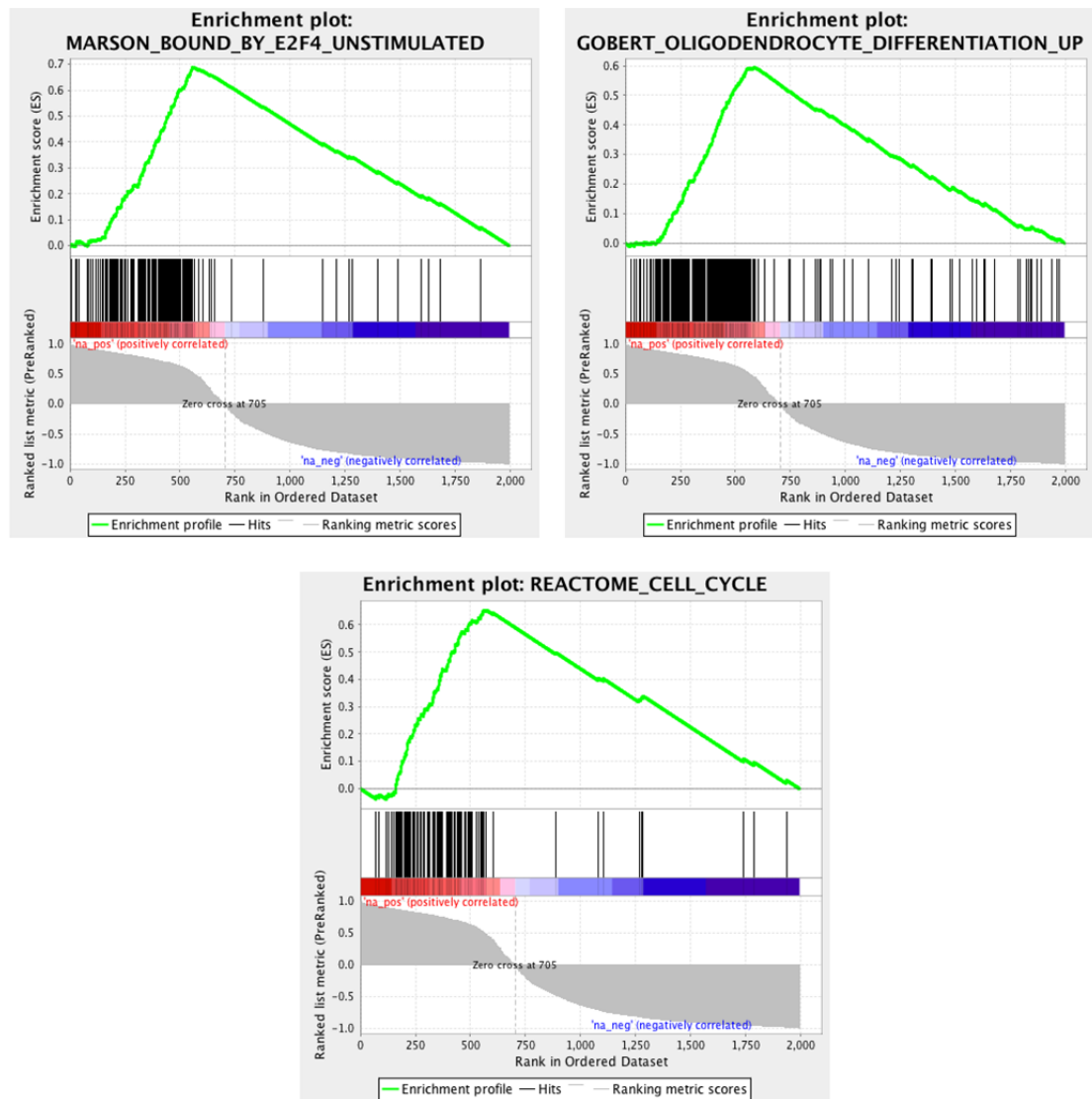


Figure 7.18 GSEA enrichment plots of significantly enriched gene sets including genes with promoters bound by E2F4, Oligodendrocyte differentiation pathway, and Reactome cell cycle annotation. These gene sets are positively enriched in empty vector controls (red) vs SMARCB1 re-expressing cells (blue). The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas the value of the ranking metric for each gene in the gene set is given by a vertical black tick.



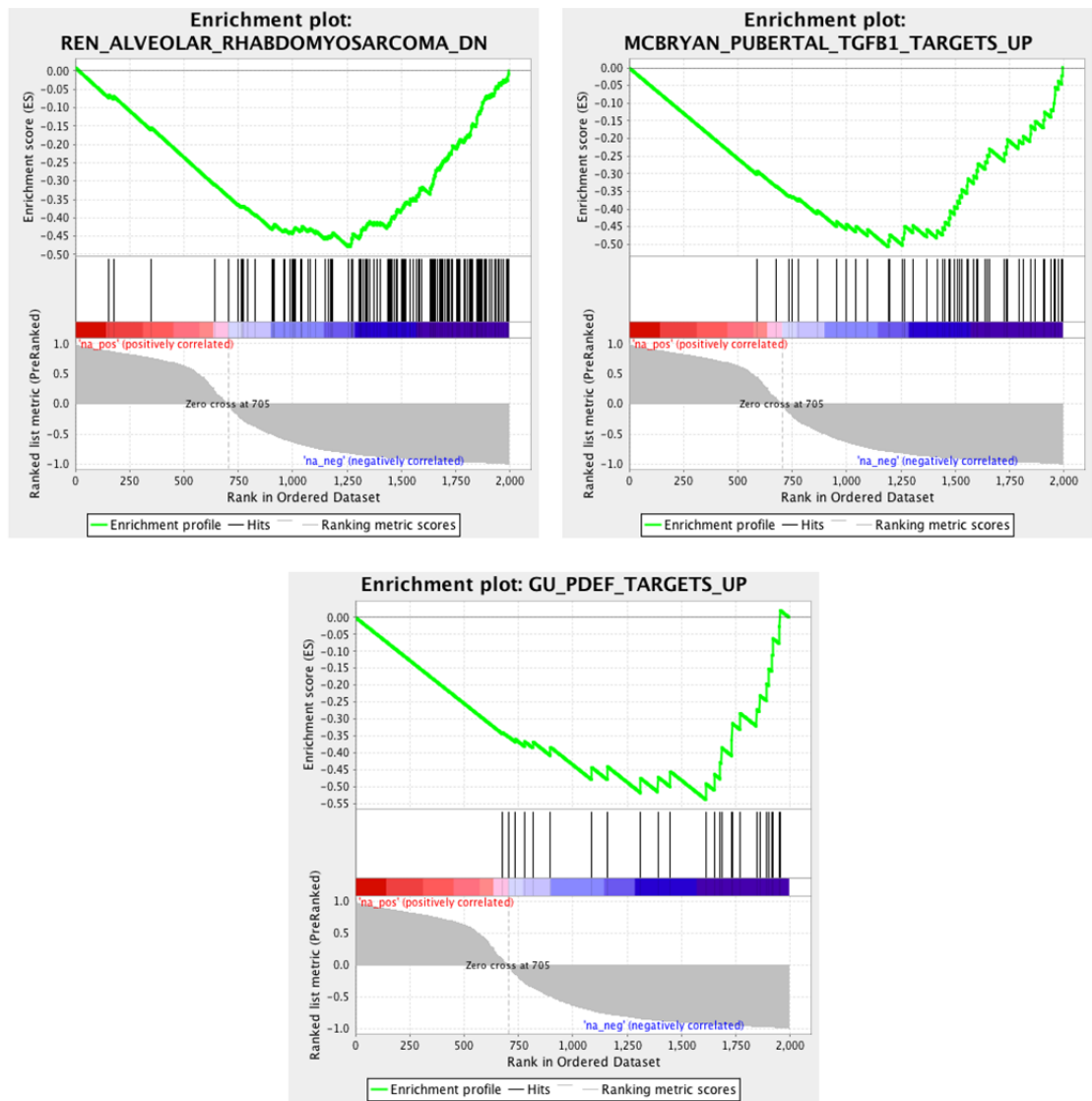


Figure 7.19 GSEA enrichment plot of significantly negatively enriched gene sets upregulated including Alveolar Rhabdomyosarcomas down-regulated genes, *TGFB1* target pathways and *PDEF* target gene sets. These gene sets are negatively enriched after *SMARCB1* re-expression (blue) vs the control infectants (red). The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas the value of the ranking metric for each gene in the gene set is given by a vertical black tick.

This is in accordance with previous findings; re-expression of *SMARCB1* in Rhabdoid cell lines causes deregulation of genes involved in the cell cycle, promoting cell differentiation, and down-regulating E2F target genes.

Analysis against the MSigDB C6 library identified 14 significantly enriched (FDR < 25% and *p*-value <1%) gene sets out of 149 and 22 significantly negatively enriched after *SMARCB1* re-expression. The analysis indicates a significant enrichment of gene sets for genes up-regulated in mice over-expressing *E2F1* with promoters bound by E2F4 and the RB pathway (Table 7.11 and Figure 7.20). Negative enrichment was observed for genes upregulated in response to *HOXA9* knock-down and for genes up-regulated by TGFB1 (Table 7.11 and Figure 7.21).

Gene set	ES	FGR	<i>p</i> -val
<b>E2F3_UP.V1_UP</b>	0.5296237	0	0
<b>RB_DN.V1_UP</b>	0.4704922	0.012	0.001
<b>E2F1_UP.V1_UP</b>	0.4188982	0	0
<b>HOXA9_DN.V1_UP</b>	-0.50132	0.001	3.30E-4
<b>TGFB_UP.V1_UP</b>	-0.427668	0.007	0.001

*Table 7.11 Selected MSigDB C6 GSEA pathways significantly enriched or negatively enriched in empty vector controls vs SMARCB1 re-expressing cell lines compared to. Relative enrichment score (ES), false discovery rate (FDR) and p-value are also displayed.*

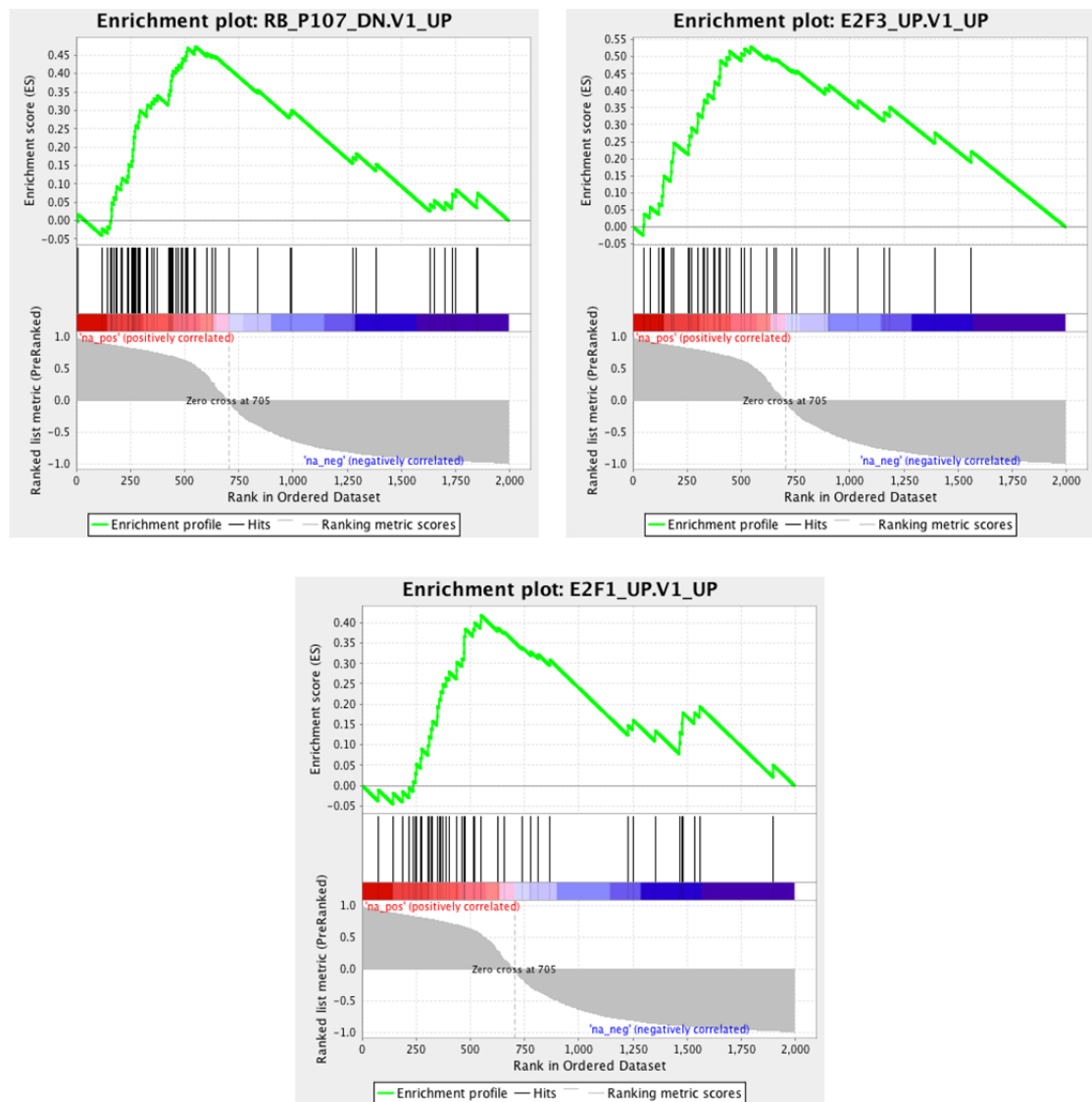


Figure 7.20 GSEA enrichment plot of genesets upregulated in response to RB and RBL1, and genesets upregulated in response to E2F3 and E2F1. These gene sets are significantly enriched in control infectants (red) vs cells re-expressing SMARCB1 (blue). The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas value of the ranking metric for each gene in the gene set is given by a vertical black tick.

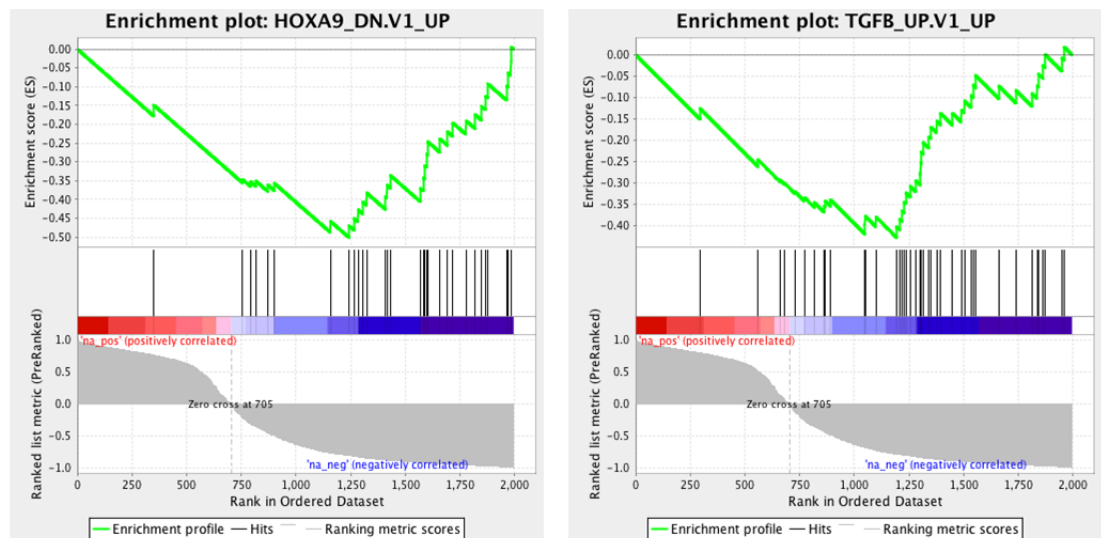


Figure 7.21 GSEA enrichment plot of gene sets related to *HOXA9*, *TGFB1* pathways. These gene sets are significantly negatively enriched after *SMARCB1* re-expression (blue) vs the control infectants (red). The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas value of the ranking metric for each gene in the gene set is given by a vertical black tick.

Analysis against MSigDb library C2 was again in accordance with previous investigations: RB not only is required in Rhabdoid cells to recruit *CDK2A*, moreover E2F1 is downstream of RB and re-expression of *SMARCB1* is well known to repress E2F1 target genes (Medjkane, Novikov et al. 2004). Also Rhabdoid Tumours show a low level of *HOXA9* and in our bioinformatics analysis the homologous protein *HOXD3* was differentially methylated. The results indicate that the TGFbeta pathway is also suppressed in Rhabdoid Tumours in a *SMARCB1* dependent manner.

#### 7.4.2.4 GSEA analysis of differentially regulated signalling pathways in Rhabdoid cell lines treated with 5-aza-2-deoxycytidine

GSEA analysis was performed by analysing a pre-ranked list of genes correlated with a metagene that defined the difference between untreated controls and 5-aza-CdR treated Rhabdoid cell lines. This pre-ranked list was analysed by GSEA against MSigDB libraries C2 and C6.

Analysis against the MSigDB C2 gene set library identified 36 significantly enriched genesets (FDR < 25% and *p*-value <1%) out of 782 and 47 significantly negatively-enriched following 5-azaCdR treatment. The significantly enriched gene sets include those corresponding to genes down-regulated in osteosarcoma upon knockdown of both *HDAC1* and *HDAC2* and genes down-regulated by TGFB1 stimulation in embryonic fibroblast and genes down-regulated in prostate cancer after knockdown of *EZH2* (Table 7.12 and Figure 7.22 ).

Gene set	ES	FDR	<i>p</i> -val
<b>SENESE_HDAC1_AND_HDAC2_TARGETS_DN</b>	0.5937639	0.001	0
<b>PLASARI_TGFB1_TARGETS_10HR_DN</b>	0.4454557	0.014	0
<b>NUYTEN_EZH2_TARGETS_DN</b>	0.3709596	0.014	0

*Table 7.12 Selected C2 GSEA gene sets significantly enriched after 5-azaCdR treatment in Rhabdoid cell lines. Relative enrichment score (ES), false discovery rate (FDR) and *p*-value are also displayed.*

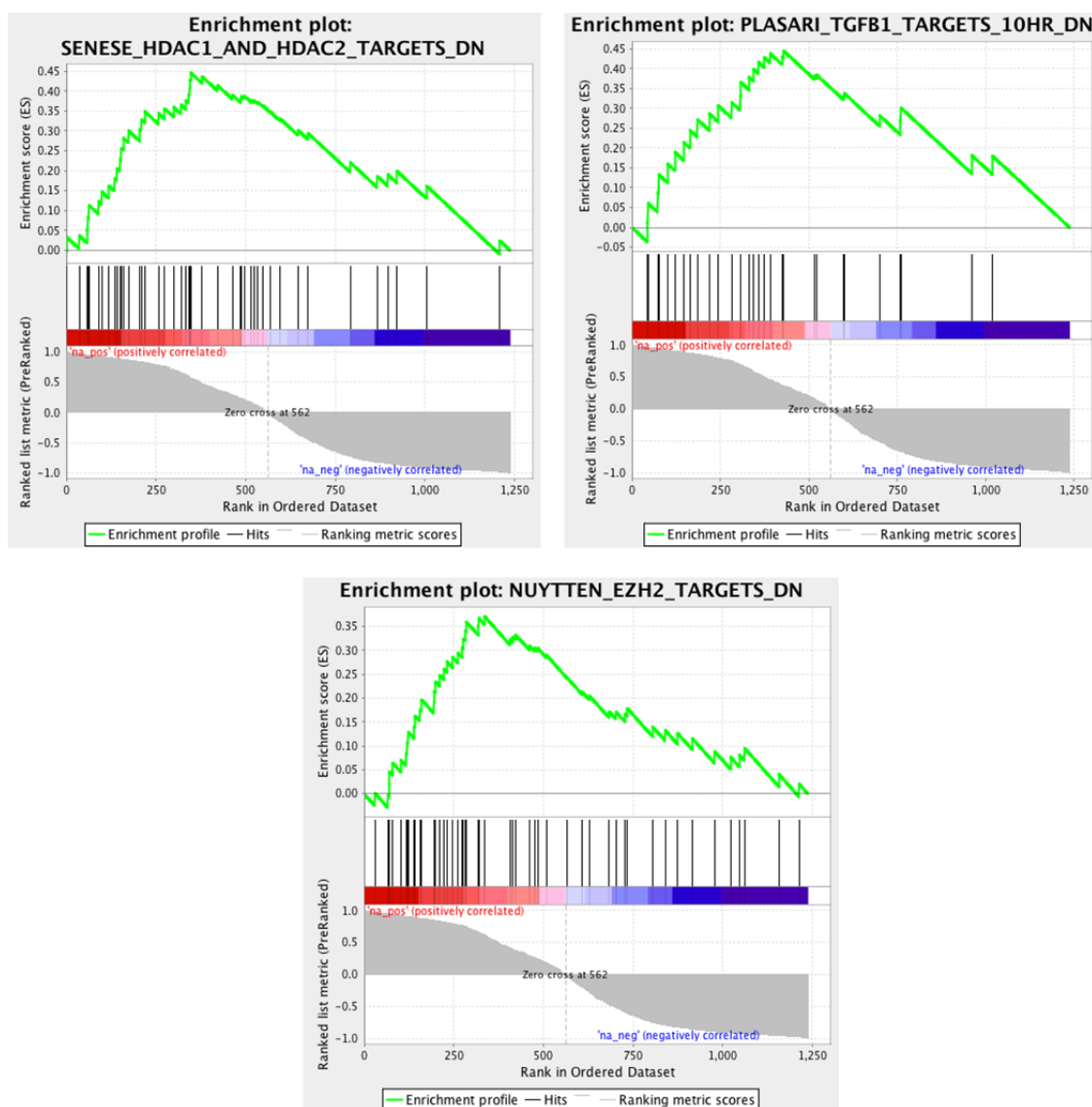


Figure 7.22 GSEA enrichment plots of selected significantly enriched genesets describing targets of HDAC1 and HDAC2, down regulated in response to TGFbeta activation and EZH2 targets. These genesets are positively enriched in Rhabdoid cell lines in the untreated cells (red) compared to cell treated with 5-azaCdR (blue). The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas value of the ranking metric for each gene in the gene set is given by a vertical black tick.

Analysis against the MSigDb C6 library identified 3 significantly enriched (FDR < 25% and  $p$ -value <1%) genesets of 115 genes set and 5 significantly negatively enriched following 5-azaCdR treatment. The analysis indicates a significant enrichment of gene sets for genes down-regulated following knockdown of *HOXA9* and a significant negative enrichment with genesets consisting of genes downregulated by use of a Notch inhibitor and with genesets corresponding to genes upregulated in response to *SMARCB1* knockout (Table 7.12 and Figure 7.23).

Gene set	ES	FGR	$p$ -val
<b>HOXA9_DN.V1_UP</b>	0.3709591	0.09	0
<b>NOTCH_DN.V1_DN</b>	-0.407226	0.22	0.016
<b>SNF5_DN.V1_UP</b>	-0.400573	0.39	0.05

*Table 7.13 Selected MSigDB c6 significantly enriched and negatively enriched gene sets following 5-azaCdR treatment in Rhabdoid cell line. Relative enrichment score (ES), false discovery rate (FDR) and  $p$ -value are also displayed.*

GSEA analysis indicates that 5-azaCdR treatment in Rhabdoid Tumours causes a similar effect of re-expression of *SMARCB1* and further suggests its overlapping function with modulation of methylation. *HDAC1* is found to be overexpressed in Atypical Teratoid Rhabdoid Tumours and the downregulation of this gene is suggested as therapeutically relevant (Sredni, Halpern et al. 2013). In addition, inhibition of *EZH2* has been indicated to cause a tumour regression in Rhabdoid malignancies (Knutson, Warholic et al. 2013) and deregulation of the Notch pathway and gene targets of *HOX9* and TGFbeta pathway highlight some of the common consequence of methylation regulation between 5-azaCdR treatment and *SMARCB1*re-expression.

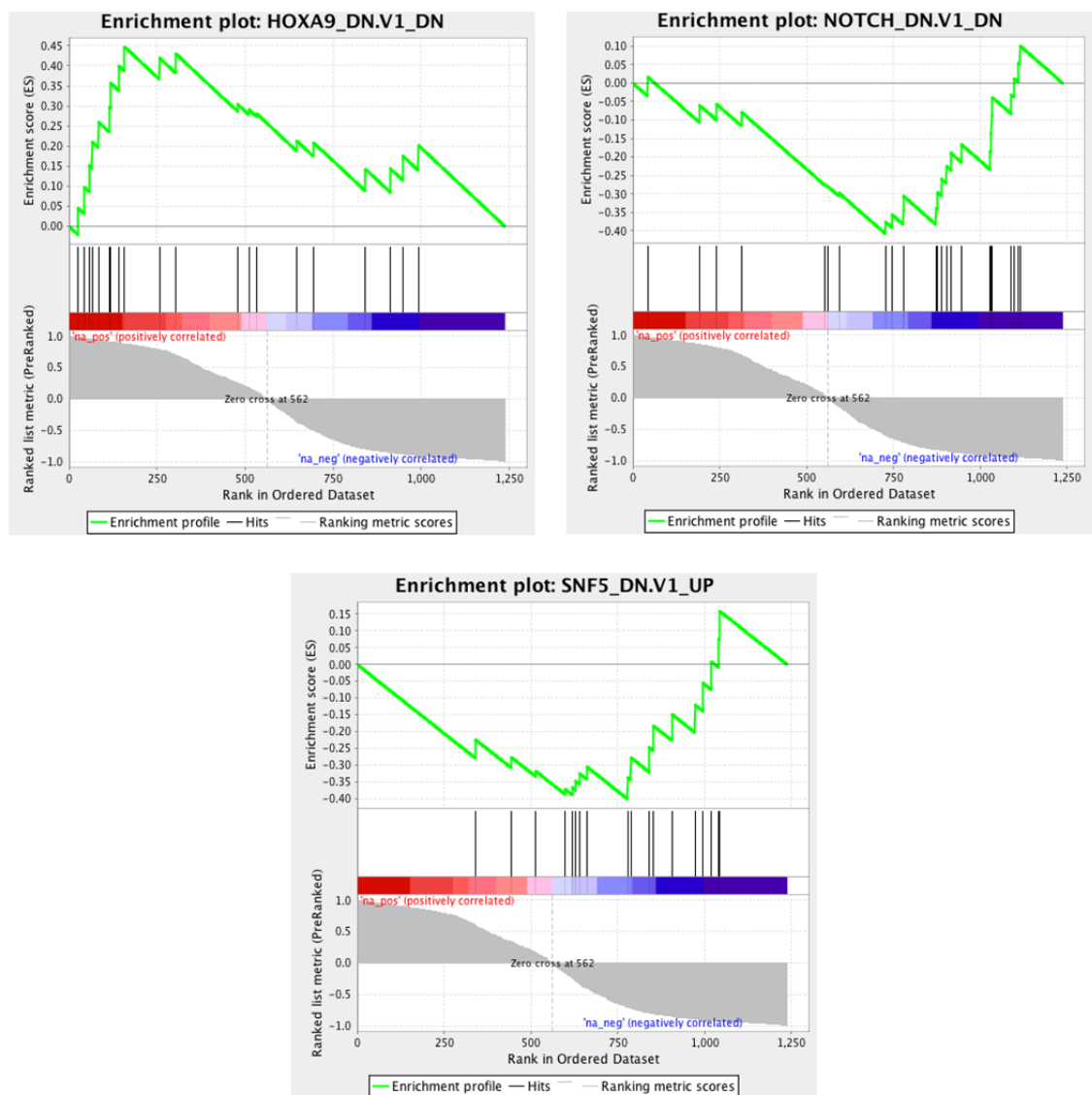


Figure 7.23 GSEA enrichment plot of selected significantly enriched genesets including targets of *HOXA9*, *NOTCH* and *SMARCB1* Knockout. These genesets are significantly enriched in untreated cells (red) vs Rhabdoid cell lines after 5-azaCdR treatment (blue). The top portion of the figure plots show the moving enrichment scores (ES) adjusted for each gene in the gene set, whereas the value of the ranking metric for each gene in the gene set is given by a vertical black tick.

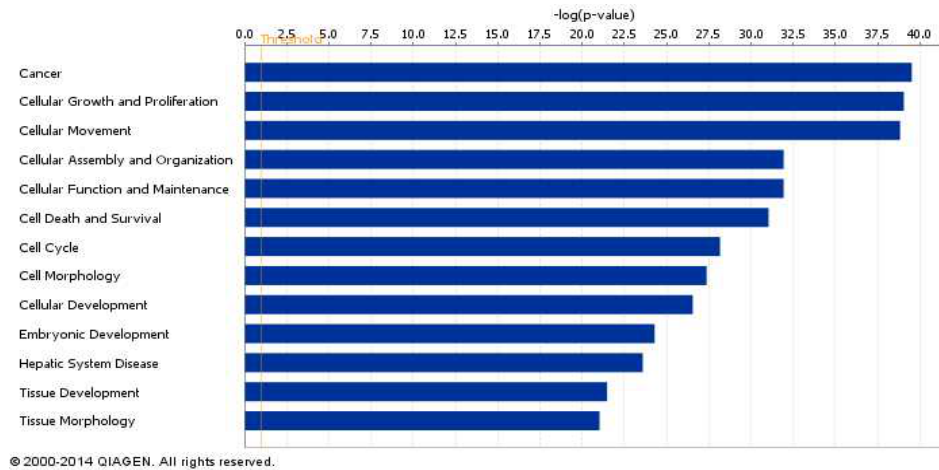


### 7.4.3 Ingenuity Pathway Analysis

In order, to identify gene expression networks which were enriched for genes highlighted in this study, Ingenuity Pathway Analysis V9 (Ingenuity®Systems, [www.ingenuity.com](http://www.ingenuity.com)) was used. Genes were ranked according to their correlation with metagenes and an adjusted p-value calculated to define when a gene was a significant part of a metagene this was defined for each metagene as previously described in section 7.4.1. Gene identifiers were mapped to corresponding genes in the Ingenuity Pathways Knowledge Base and the significance of the association between significant gene lists and each pathway was computed using a hypergeometric statistical test, calculating a *p*-value to identify association between gene lists and canonical pathways.

#### *7.4.3.1 Analysis of pathway deregulation in SMARCB1 re-expressing Rhabdoid Tumour cells*

Analysis of the genes which form a significant part of a metagene defining the expression difference between Rhabdoid cells re-expressing *SMARCB1* and empty vector controls provided the confirmation that *SMARCB1* modulates a wide range of biological functions. The genes were ranked by the extent of metagene correlation, Determining the enrichment of gene sets for each biological function. The statistical analysis determined the p-value for each functional category, identifying categories such as Cancer, Cell Death, Inflammatory Disease, Cell Death, Cellular Growth and Proliferation, Cell-Cell-Signalling, Cell Cycle and Embryonic Development (Figure 7.24).



*Figure 7.24 Important biological functions associated with genes that defined SMARCB1 re-expressing cell lines. Biological functions associated with genes enriched in a metagene defining empty vector controls vs SMARCB1 expressing Rhabdoid cells are indicated on the y-axis. Each bar corresponds to  $-\log(p\text{-value})$  for the enrichment of genes in the selected pathway. The orange vertical line represents the threshold for significance.*

IPA analysis of the data set also identified several important canonical pathways that are deregulated as a result of *SMARCB1* re-expression. These include, integrin signalling, IL-B signalling and signalling of Rho family GTPases. Interestingly, some of the functional categories highlighted were also identified in the GSEA analysis (Figure 7.25).

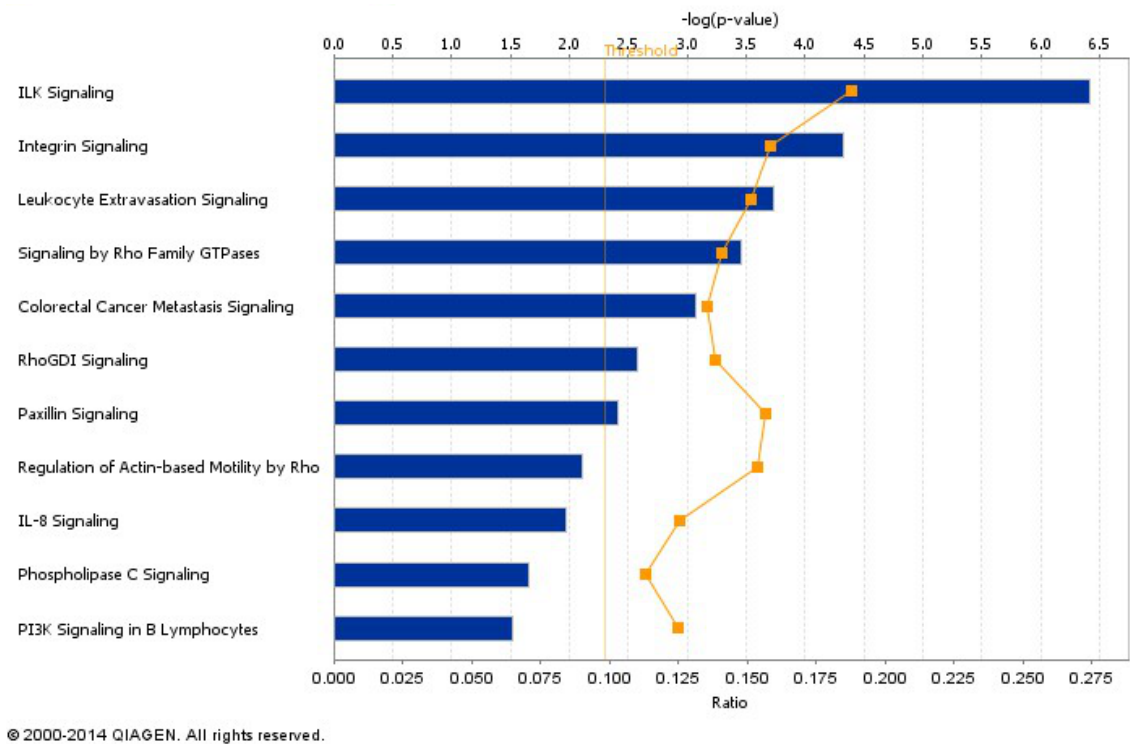
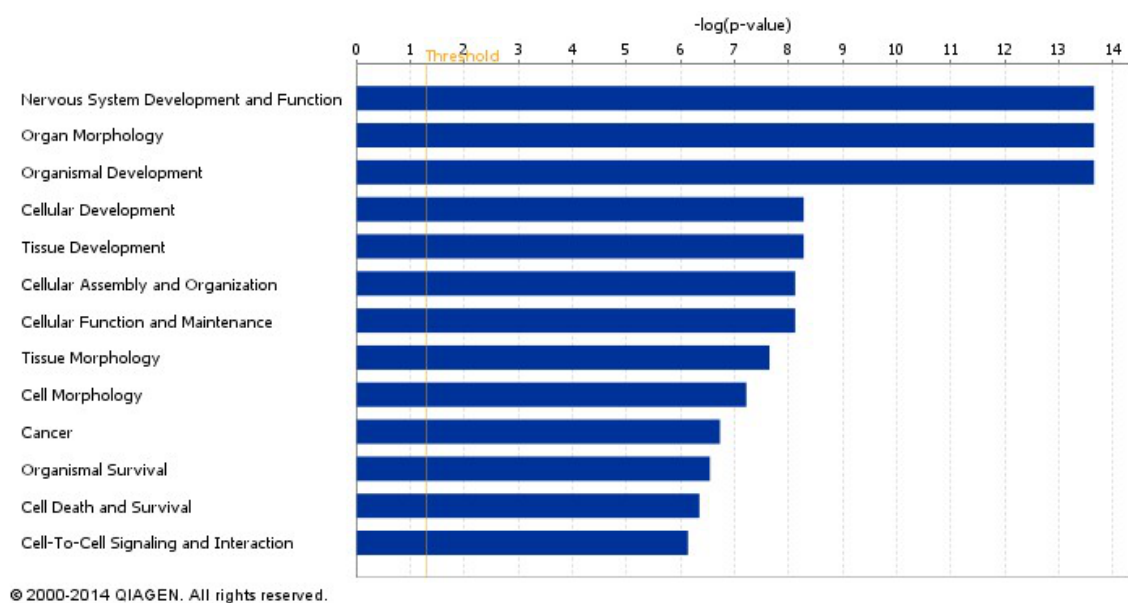


Figure 7.25 Selected canonical pathways associated with genes that defined SMARCB1 re-expressing cell lines. Biological functions associated with genes enriched in a metagene defining empty vector controls vs SMARCB1 expressing Rhabdoid cells are indicated on the y-axis. Each bar corresponds to  $-\log(p\text{-value})$  for the enrichment of genes in the selected pathway. The orange vertical line represents the threshold for significance.

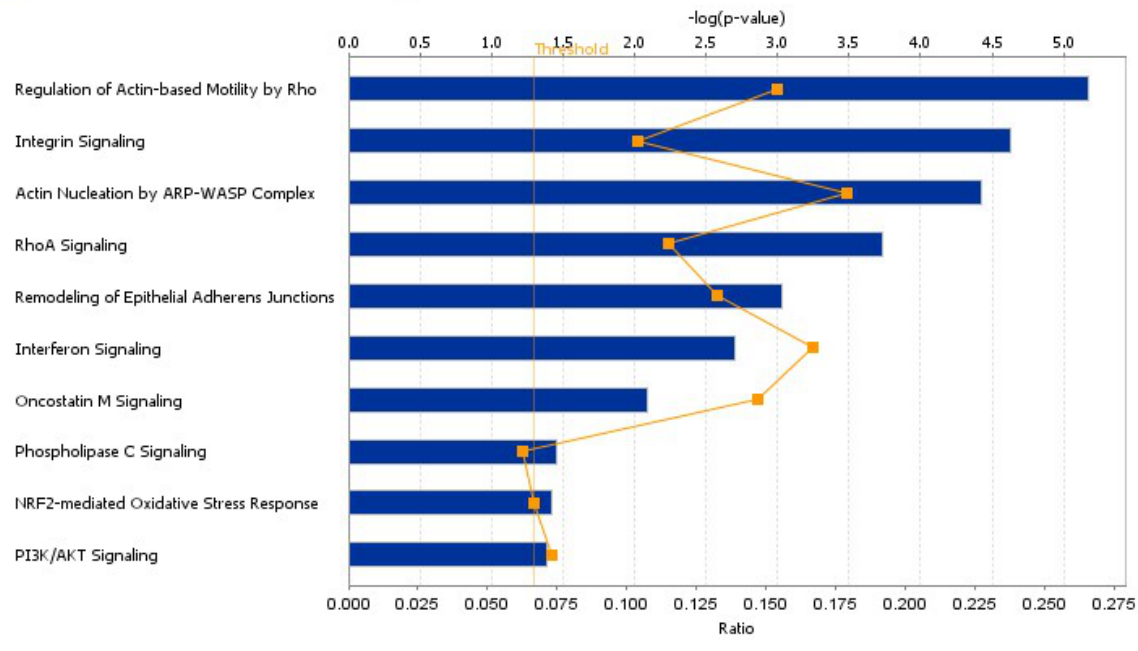
#### 7.4.3.2 Analysis of pathway deregulation in Rhabdoid Tumours

Analysis of the genes which form a significant part of a metagene defining the expression difference between primary Rhabdoid Tumours and Medulloblastoma indicate deregulation of a large number of biological functions. The genes were ranked by the extent of metagene correlation, Determining the enrichment statistic for each biological function. The analysis highlights the modulation of pathway involved in neuron system development, organ and cell morphology and cell development (Figure 7.26).



*Figure 7.26 Important biological functions associated with genes whose expression defines primary Rhabdoid Tumours. Biological functions associated with genes enriched in a metagene defining the difference in expression between primary Rhabdoid Tumours and Medulloblastoma are indicated on the y-axis. Each bar corresponds to  $-\log(p\text{-value})$  for the enrichment of genes in the selected pathway. The orange vertical line represents the threshold for significance.*

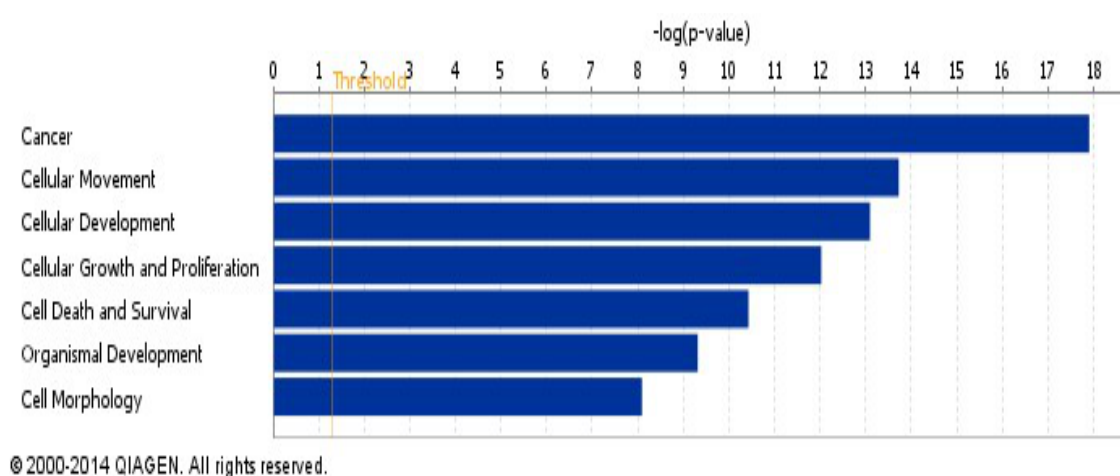
Analysis of the enrichment of canonical pathways identifies deregulation of several pathways including Rho signalling, integrin signalling and P13K/AKT signalling (Figure 7.27).



*Figure 7.27 Selected canonical pathways associated with genes whose expression defines Rhabdoid Primary tumours. Biological functions associated with genes enriched in a metagene defining the difference in expression between primary Rhabdoid Tumours and Medulloblastoma are indicated on the y-axis. Each bar corresponds to  $-\log(p\text{-value})$  for the enrichment of genes in the selected pathway. The orange vertical line represents the threshold for significance.*

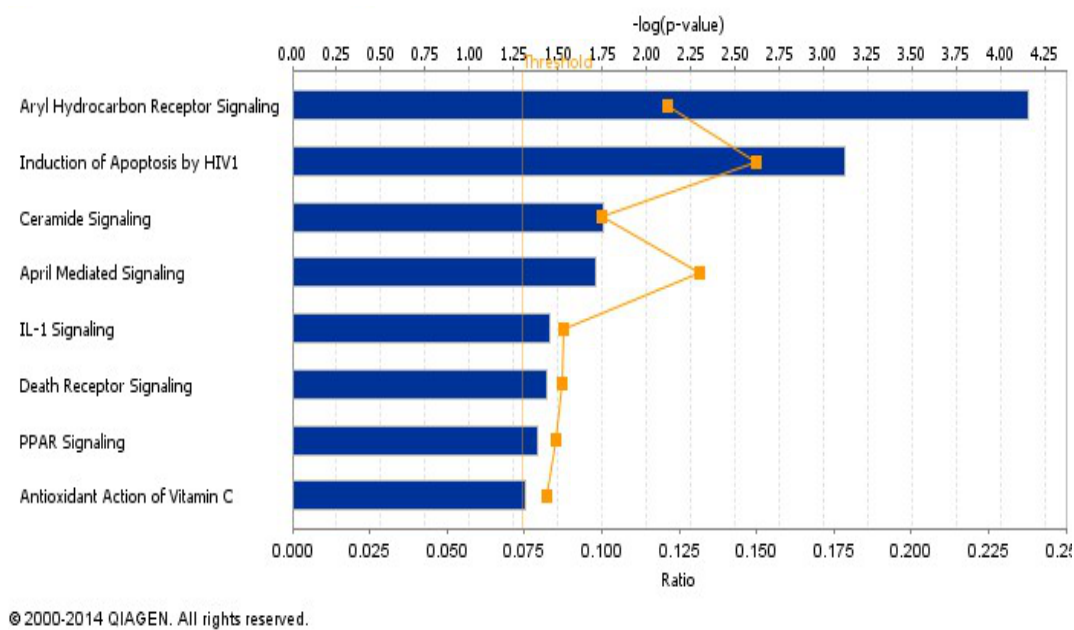
#### 7.4.3.3 Analysis of pathway deregulation in 5-aza-CdR treated Rhabdoid Tumour cells

Analysis of the genes which form a significant part of a metagene defining the expression difference between untreated cells and Rhabdoid cells treated with 5-aza-2-deoxycytidine indicate significant deregulation of a number of biological functions involved in cancer cellular movement and development, cell death and morphology. The genes were ranked by the extent of metagene correlation, Determining the enrichment statistic for each biological function (Figure 7.28).



*Figure 7.28 Selected significant biological functions associated with 5-azaCdR treatment in Rhabdoid Tumours. Biological functions associated with genes enriched in a metagene defining the difference in expression between untreated cells and 5-aza-CdR treated Rhabdoid Tumour cells are indicated on the y-axis. Each bar corresponds to  $-\log(p\text{-value})$  for the enrichment of genes in the selected pathway. The orange vertical line represents the threshold for significance.*

Analysis of the enrichment of canonical pathways highlights pathways involved in apoptosis, IL1 signalling, death receptor and PPAR signalling (Figure 7.29).



*Figure 7.29 Selected canonical pathways associated with 5-azaCdR treatment in Rhabdoid Tumours. Biological functions associated with genes enriched in a metagene defining the difference in expression between untreated cells and 5-aza-CdR treated Rhabdoid Tumour cells are indicated on the y-axis. Each bar corresponds to  $-\log(p\text{-value})$  for the enrichment of genes in the selected pathway. The orange vertical line represents the threshold for significance*

#### 7.4.3.4 Custom pathway analysis

*SMARCB1* re-expression and 5-aza-2-deoxycytidine treatment in Rhabdoid cell line results in constitutive activation of transcription factors associated with cell death signalling, differentiation signalling and with control of cell cycle progression. To investigate the outcome of these modulations, the metagene content derived from the model experiment and the primary analysis were further investigated by direct comparison. The comparison was performed based on the gene ranking of each data set assigned by IPA ingenuity.

Comparative analysis of the pathways involved identified a set of canonical pathway that were significantly down-regulated or upregulated in both *SMARCB1* re-expressing and 5-azaCdR treated Rhabdoid cells with a corresponding significant effect in Rhabdoid primary malignancies. In particular, the analysis predicted a significant inactivation of integrin signalling and downregulation of actin-based motility of Rho signalling by Rho family GTPases pathways in response to *SMARCB1* re-expression and 5-azaCdR treatment, and predicted relative activation in the primary malignancies (Figure 7.30).

A similar strategy was also used to identify significant upstream regulatory effectors that are common to all data sets. Upstream regulatory effectors are identified by IPA based upon a statistic describing the ability of an individual regulator to produce the expression pattern observed. The more of the deregulated expression that can be potentially explained by an upstream effector the greater the p-value. This computational analysis was performed to indicate which biological processes are affected by *SMARCB1* re-expression. The investigation identified a variety of effectors, which were predicted to be activated in response to *SMARCB1* re-expression and 5-azaCdR treatment and oppositely downregulated in Rhabdoid primary tumours. Interestingly, one of the most significant effectors was TGFB1, a pathway whose involvement was also determined by GSEA analysis (Figure 7.31).

Molecular activity prediction analysis was used to predict and hypothesise a plausible pathway, based upon IPA's custom pathways feature and evidence of significant



enrichment of genes within the data sets. Specifically, the TGF-beta pathway was further investigated. As indicated in Figure 7.32, Figure 7.33 and Figure 7.34 a significant enrichment of targets regulated by TGFbeta receptor were identified across the data set: in particular *SMARCB1* and 5-azaCdR treatment positively regulated the pathway in a similar manner, while in Rhabdoid primary data the same pathway was deregulated in an opposite manner.

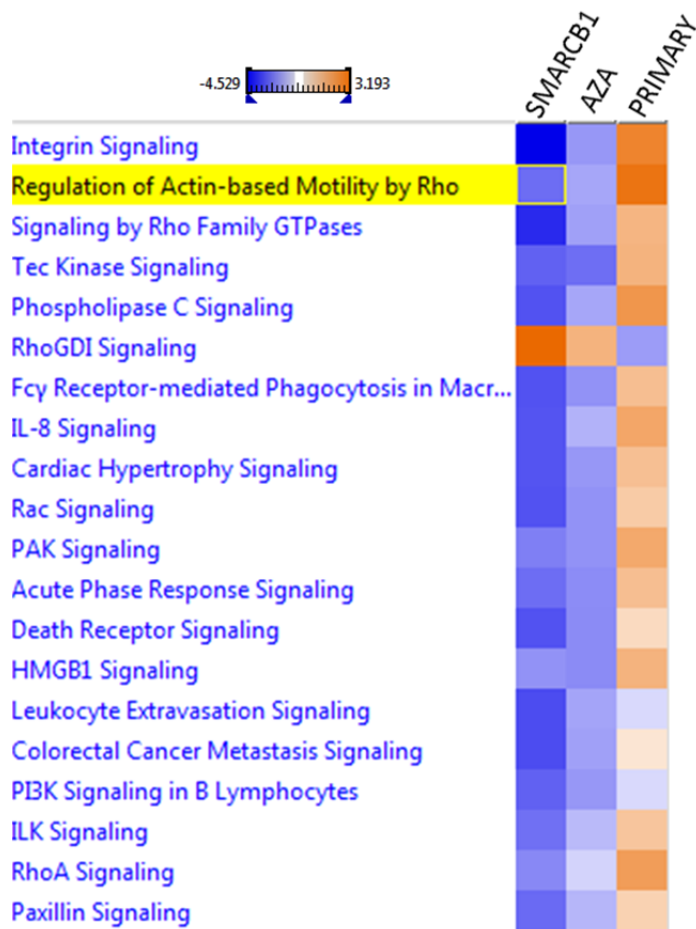


Figure 7.30 Heatmap of the top 20 common canonical pathways identified by comparative analysis. Intensity represents standard deviations from the mean (Z-score). Inactivation (pale blue to dark blue) and activation of pathway (pale red to dark red) is represented by the appropriate colour. Each row represents a canonical pathway and each column a data set. *SMARCB1* and *AZA* refer to Rhabdoid cells expressing *SMARCB1* or treated with 5-AZA-CdR. Primary refers to primary Rhabdoid Tumours.

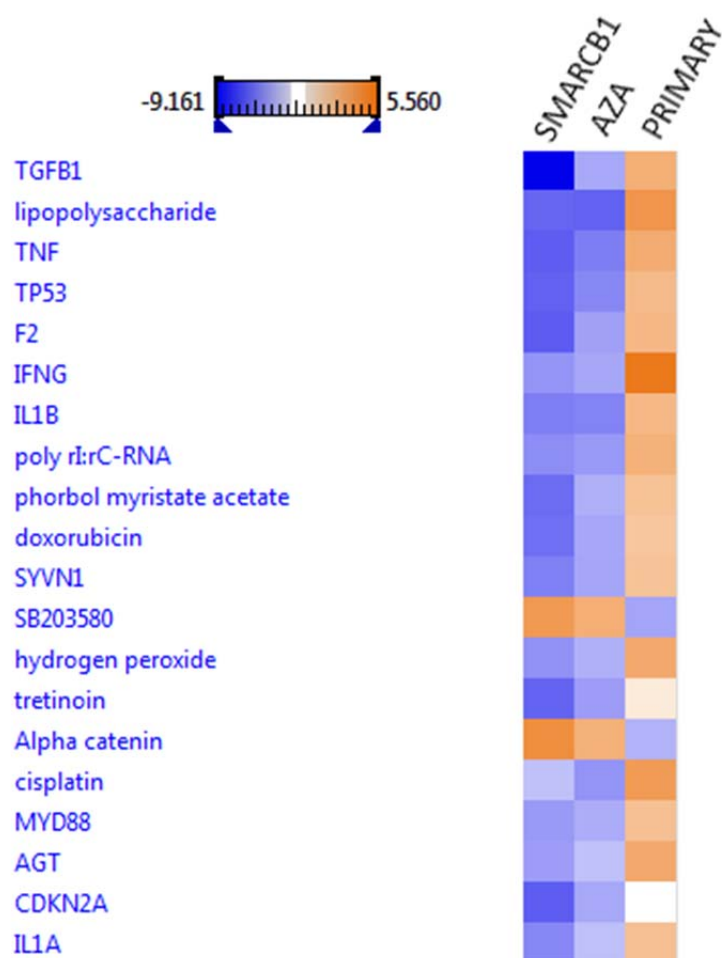
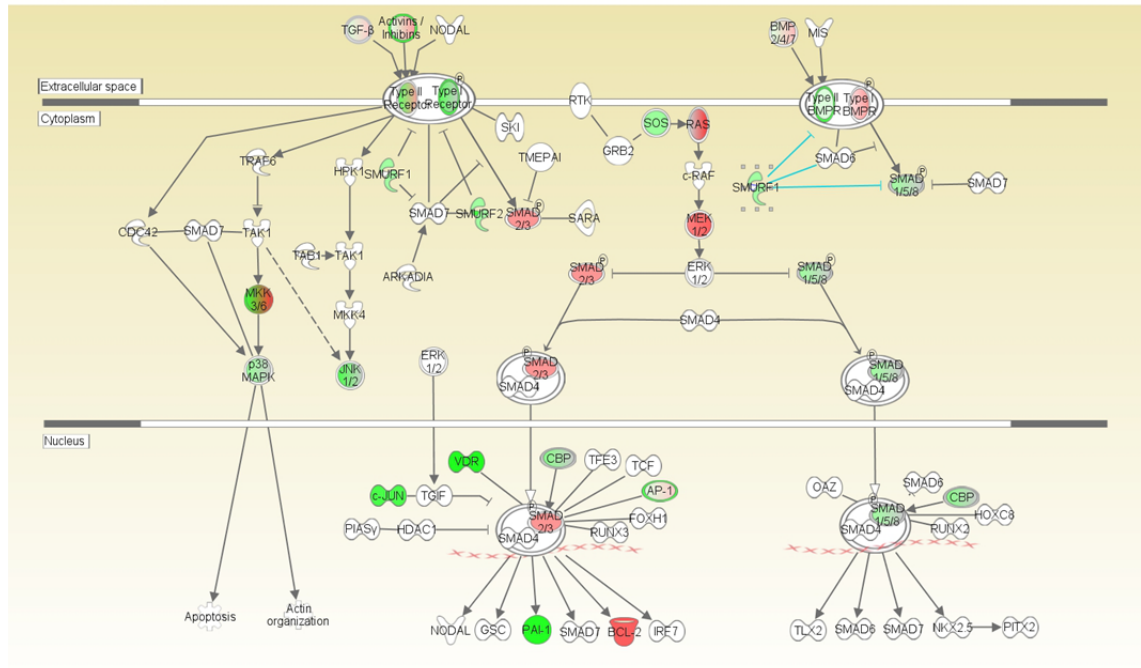
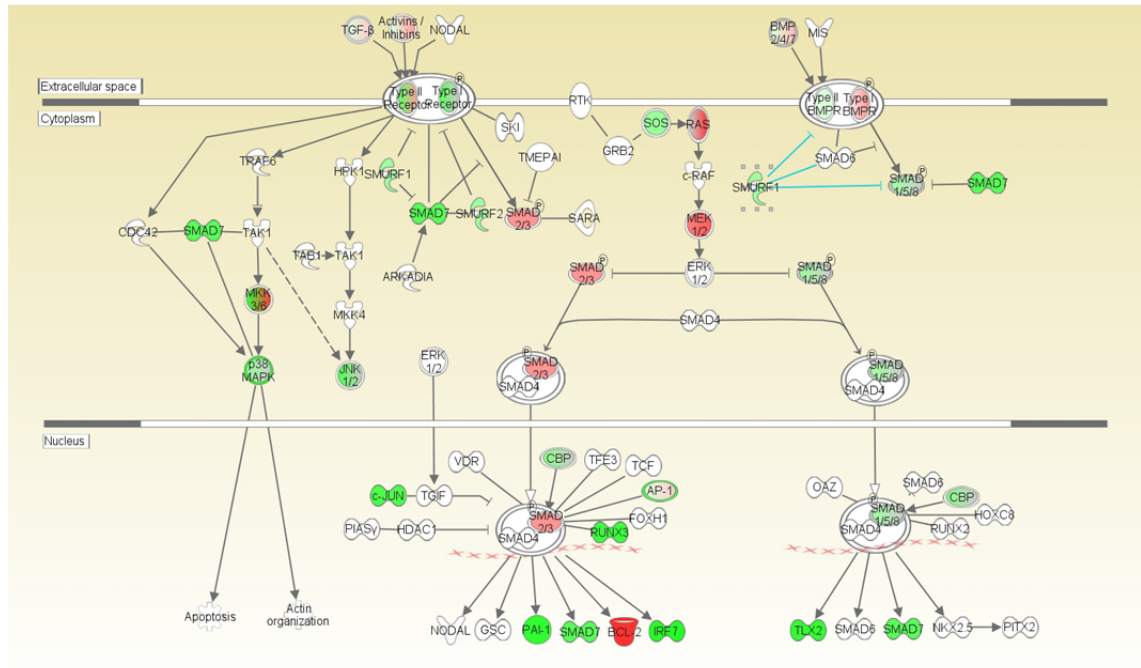


Figure 7.31 Heatmap of the top 20 upstream effectors identified by comparative analysis. Intensity represents standard deviations from the mean (Z-score). Activation (pale blue to dark blue) and inactivation of pathway (pale red to dark red) is represented by the appropriate colour. Each row represents a canonical pathway and each column a data set. SMARCB1 and AZA refer to Rhabdoid cells expressing SMARCB1 or treated with 5-AZA-CdR. Primary refers to primary Rhabdoid Tumours



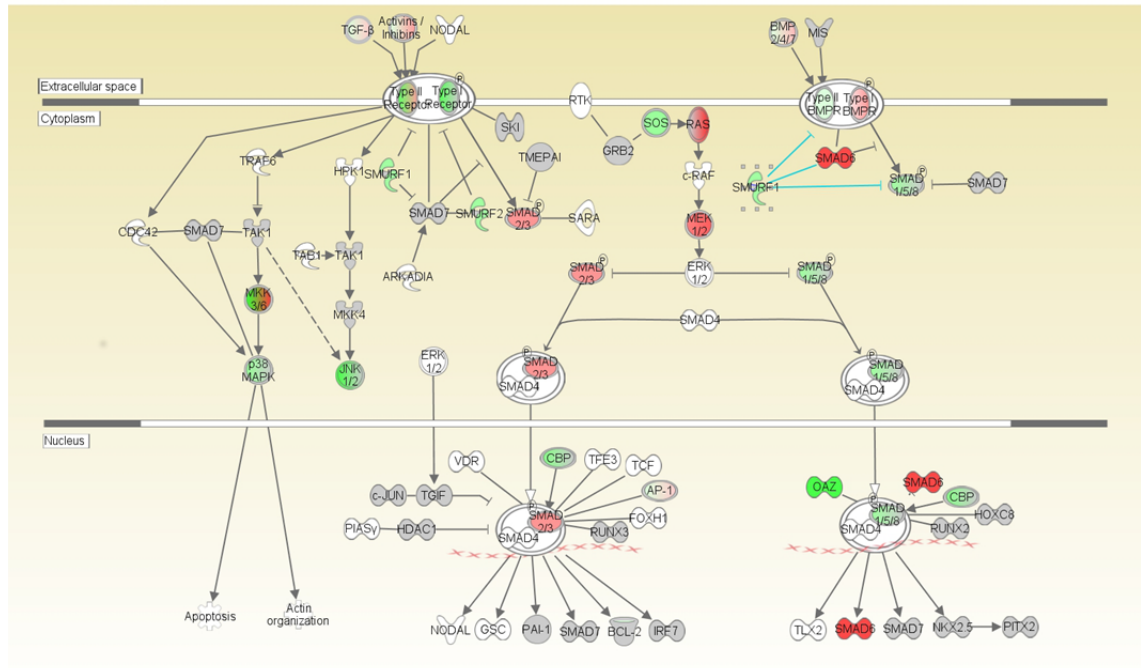
© 2000-2014 QIAGEN. All rights reserved.

Figure 7.32 Modulation of the TGF-beta pathway is associated with Rhabdoid Tumourigenesis. Expression data from SMARCB1 re-expressing cells was overlaid onto the canonical pathway using IPA. The genes which are coloured in bold red and green are genes deregulated in Rhabdoid primary tumours (red = up, green = down). Green and red arrows indicate a predicted down and up-regulating interaction respectively.



© 2000-2014 QIAGEN. All rights reserved.

Figure 7.33 Modulation of the TGF-beta pathway is associated with Rhabdoid Tumourigenesis. Expression data from 5-azaCdR treated cells was overlaid onto the canonical pathway using IPA. The genes which are coloured in bold red and green are genes deregulated in Rhabdoid primary tumours (red =up, green = down). Green and red arrows indicate a predicted down and up-regulating interaction respectively.



© 2000-2014 QIAGEN. All rights reserved.

Figure 7.34 Modulation of the TGF-beta pathway is associated with Rhabdoid Tumourigenesis. Expression data from Rhabdoid primary samples was overlaid onto the canonical pathway using IPA. The genes which are coloured in bold red and green are genes deregulated in Rhabdoid primary tumours (red =up, green = down). Green and red arrows indicate a predicted down and up-regulating interaction respectively.

## 7.5 Conclusion

### 7.5.1 Role of *SMARCB1* in methylation modulation

An investigation of the role of *SMARCB1* in epigenetic modulation was performed through a systematic analysis of the data across different platform and experimental conditions. Firstly changes in methylation status following *SMARCB1* re-expression was compared with 450K methylation array data from primary tumours. The analysis identified 63 CpGs that were significantly demethylated in response to *SMARCB1* and normally methylated in Rhabdoid primary tumours. Bisulfite Sequencing of *ABR* (regulator of the Rho family), *BHLHE40* (transcription repressor that induces apoptosis) and *GALNT2* (promoter of tumour invasion) confirmed the hypomethylating effect of *SMARCB1* re-expression in Rhabdoid cell lines.

To better understand the mechanism by which *SMARCB1* induces hypomethylation in Rhabdoid cell lines, significant expression data from cell lines re-expressing *SMARCB1* and treated with the demethylating agent 5-azaCdR were cross-referenced. The analysis identified 14 genes deregulated in both conditions. The list contains three genes that were also found differentially expressed following *SMARCB1* re-expression (*DNA2*), in response to 5-azaCdR treatment (*PDE4C*- *PDE4D*) and in primary tumours (*SERPINE2*). Also to understand the effect of hypomethylation on gene expression caused by *SMARCB1*, cross-referencing of RNAseq and 450K platforms in *SMARCB1* re-expressing data was performed. A selection of genes *GLI2* and *GLI3* (Shh pathway regulator), *JARID2* (promoter of undifferentiated Rhabdomyosarcomas) and *NTN4* (inhibitor of angiogenesis) were further analysed. In particular qPCR of RNA from *SMARCB1* re-expressing Rhabdoid cells and 5-azaCdR treated Rhabdoid cells revealed an overexpression of these genes to some extent in almost all genes in both conditions but most notably in *GLI2*. This validated the hypothesis that *SMARCB1* modulates gene expression in part by provoking a methylation modification in Rhabdoid Tumours.

## 7.5.2 Pathway analysis

Analysis of the pathways involved in Rhabdoid tumourigenesis was performed using a different method. Overall cross-referencing across expression data sets was performed in order to disclose gene signature and pathways involved in Rhabdoid tumourigenesis.

### 7.5.2.1 *Gene expression cross-referencing identifies signature of Rhabdoid phenotype*

First, gene expression data set cross-referencing was applied across the experimental conditions and primary tumours aimed at exploring gene signatures across data sets that could defined a Rhabdoid phenotype and characterise direct targets of *SMARCB1* loss. In particular, the gene content of appropriate metagenes was overlapped, to identify 13 genes *SMARCB1*/methylation dependent genes correlated with a Rhabdoid phenotype. Interestingly, all 13 genes are involved in tumourigenesis in a variety of cancer though for most the mechanisms and exact function is not fully understood e.g. *CCDC74B* (implicated in mitosis regulation and cell cycle regulation), *Dnm3os* ( gene transcribed into a non-coding RNA indispensable for body growth in mammals), *IDH1* (mutation of this gene has an oncogenic function in Gliomas), *KAZALD1* (highly methylated in Sarcomatoid-type peritoneal mesothelioma) and *KIAA0101* (overexpressed in breast cancer and modifies directly centrosomes).

### 7.5.2.2 *GSEA analysis identifies deregulated pathways responsible for tumourigenesis in Rhabdoid Tumours*

Metagene contents were analysed using GSEA to classify gene sets that share common biological (C2: cured annotations) and tumourigenic functions (C6: oncogenic signature).

GSEA analysis of primary Rhabdoid Tumour metagene content indicated a variety of pathways deregulated. In particular, significant enrichments in Rhabdoid patients were observed for genes commonly down-regulated in human alveolar rhabdomyosarcomas when compared with Medulloblastomas; genes up-regulated in embryonic stem cells and genes involved in G1/S DNA damage checkpoint in a p53 independent manner.

Significant negative enrichment instead was detected for gene sets that resemble the Medulloblastomas phenotype. This was in accordance with what was previously reported in the literature and in this study.

Study of the metagene that defined Extra Cranial Rhabdoid Tumours (ECRT) gave, not surprisingly, fewer significantly enriched gene sets correspond to deregulated pathways. A significant enrichment of gene sets describing downregulation of targets of *PATCH1* and *SUFU* was found; and a significant negative enrichment of a gene set corresponding to genes up-regulated in kidney epithelium cells over-expressing *KRAS* and to genes up-regulated in breast cancer cells over-expressing *YAP*.

Re-expression of *SMARCB1* in Rhabdoid Tumours significantly deregulates a wide number of pathways. Most significantly enriched genesets reflect in part the changes in phenotype that occurs in Rhabdoid cell upon *SMARCB1* re-expression i.e. cell cycle progression and recruitment of SWI/SNF complexes. Significant enrichment was detected for gene sets involved in cell cycle and differentiation, and corresponding to genes with promoters bound by *E2F4*. Significant negative enrichment was found for alveolar Rhabdomyosarcomas gene set pathway, the TGF-Beta target pathway and downregulation in response to *HOXA9* knockdown and *PDEF* target.

GSEA analysis was also performed on metagenes that defined Rhabdoid cells treated with 5-azaCdR: significant enrichment of gene sets describing genes down-regulated following knockdown of *HOXA9* and a significant negative enrichment of a gene set composed of genes downregulated when using a Notch inhibitor and a gene set corresponding to genes upregulated in response of *SMARCB1* knockout.

In conclusion, the GSEA analysis underlines the differences between the data set previously described in this study. Also it reconfirmed the subtle differences between Atypical Teratoid Rhabdoid Tumours and in Extra Cranial Rhabdoid Tumours and the similar effect on gene expression caused by *SMARCB1* re-expression and 5-azaCdR treatment in Rhabdoid cells.



### 7.5.2.3 Ingenuity Pathway analysis

IPA pathway analysis was performed in order to characterise a common pathway deregulated across functional models and primary data.

Expression data were analysed separately, confirming what was observed by GSEA analysis. Specifically, a general deregulation of cell cycle gene control, cell morphology and cell developmental pathways were observed across all the data sets, as expected. Interestingly, canonical pathway analysis reveals deregulation of integrin pathway and Rho-related pathway across both functional models and primary data. Comparative analysis of the data sets in IPA identified 21 canonical pathways that were upregulated in response to *SMARCB1* re-expression and 5-azaCdR and downregulated in the Primary tumours. Most of these pathways were also observed in GSEA, such as Signalling of Rho family GTPse, Glioblastoma pathways, etc. Moreover, comparative analysis identifies 13 significant upstream regulators, included TGFB1, TNF, TP53 and IFNG. The TGFbeta pathway was further analysed and molecular activity prediction was computed for this pathway.

The cross-referencing data analysis indicated that TGFbeta was deregulated negatively in primary Rhabdoid Tumours , while in *SMARCB1* re-expressing cells and 5-azaCdR treated cells an opposite effect was observed. TGFbeta is a potent growth inhibitor and perturbation of the signalling pathway promotes cancer cell growth and finally metastasis (Akhurst and Derynck 2001). Importantly TGFbeta is linked with other effectors deregulated in our study, such as the Rho family (*ABR*), and GLI2 (early gene target of the TGF- $\beta$ /SMAD pathway) (Dennler, Andre et al. 2007).

## **Chapter 8      Conclusion**

## 8.1 Introduction

Malignant Rhabdoid Tumours are an aggressive malignancy that occurs in infants and young children. They can occur in soft tissue (Extra Cranial Rhabdoid Tumours ), in the brain and in the spinal cord (Atypical Teratoid Rhabdoid Tumours ). This malignancy presents a biallelic inactivation at chromosome 22 of *SMARCB1*, this deletion characterises these tumours (Jackson, Sievert et al. 2009). However, malignancies located in the brain can be misdiagnosed as Medulloblastoma or PNET, since they share locations and histological features. Moreover, these malignancies are treated with combination of therapies including surgery, radiation and high dose chemotherapy, often producing severe side effects (EURHAB).

*SMARCB1*'s role in tumorigenesis has been previously investigated *in vivo* and *in vitro*. *SMARCB1* +/- and conditional *SMARCB1* knock-out mice develop malignancies and undifferentiated sarcomas, indicating the oncogenic effect of *SMARCB1* loss (Bultman, Gebuhr et al. 2000, Roberts, Leroux et al. 2002). In Rhabdoid cell lines re-expression of *SMARCB1* induces arrest of the cell cycle in G1 and changes in cell morphology associated with senescence, down-regulation of *E2F* targets, activation of Rb and up-regulation of *CDKN2A* and *CDKN1A* (Versteeg, Medjkane et al. 2002, Chai, Charboneau et al. 2005). Moreover ChIP experiments demonstrate that *SMARCB1* re-expression directly recruits SWI/SNF complexes at promoter region of *CDKN2A* and *CDKN1A* and evicts the polycomb complex silencing proteins followed by H3K4 methylation and H3k27 demethylation (Kusafuka, Miao et al. 2004, Kuwahara, Charboneau et al. 2010, Kuwahara, Mora-Blanco et al. 2013). Microarray analysis in Rhabdoid cell lines identifies up-regulation of genes controlling cell cycle, senescence markers and repression of mitotic genes following *SMARCB1* re-expression (Birks, Donson et al. 2013).

*SMARCB1*'s role in tumorigenesis has been also investigated in primary Rhabdoid Tumours by whole genome profiling and Single Nucleotide Polymorphism Profiling (SNP) assays (Jackson, Sievert et al. 2009). In particular, comparison between Atypical Teratoid Rhabdoid Tumours and other CNS paediatric tumour (such as

Medulloblastoma and Glioblastoma) identified downregulation of neuronal development and differentiation genes, up-regulation of histone deacetylase inhibitors genes (HDACi) and up-regulation of *BMP* pathway genes (Ma, Yang et al. 2014). Just one study attempted to analyse the difference in expression between Atypical Teratoid Rhabdoid Tumours and in Extra Cranial Rhabdoid Tumours. Specifically, this investigation showed that in Extra Cranial Rhabdoid and Atypical Teratoid Rhabdoid Tumours share a common dysregulation of cell cycle and epigenetic genes, suggesting the possibility of a therapy effective in both malignancies (Birks, Donson et al. 2013). While gene expression profiling in Rhabdoid cell lines and Primary malignancies was performed, an analytical and comprehensive analysis of the epigenetic effects of *SMARCB1* loss has yet to be attempted. Investigating the implication of *SMARCB1* in the epigenetic machinery may open the way not only to a complete understanding of the *SMARCB1* mechanism of tumourigenesis, but also to characterising potential therapeutic targets.

The current study was conceived to provide new understanding of the tumourigenic mechanism of Rhabdoid Tumours. Though comprehensive gene expression and methylation profiling of Rhabdoid primary malignancies and of *SMARCB1* re-expressing cell lines, this research aimed to catalogue events dependent on *SMARCB1*. The data collected were finally analysed in such a way as to characterise pathways and genes associated with tumourigenesis and potentially therapeutic targets.

## **8.2 DNA methylation and gene expression profiling of Rhabdoid Tumours characterise a unique a coherent tumour type when compared with other embryonal and neuronal malignancies**

A better understanding of the molecular mechanisms driving tumourigenesis in Rhabdoid malignancies is necessary to characterise deregulated genes and possible therapeutic targets. Few studies have tried to compare gene expression of Rhabdoid primary tumours in order to disclose similarities and differences among Rhabdoid malignancies from differently locations. In particular Birks compared microarray gene expression profiles of Atypical Teratoid Rhabdoid Tumours tumour samples with already published microarray data for Rhabdoid kidney tumours. In this study found only a few common gene signatures consistently affected by *SMARCB1* expression, such as *LMO4*, *SLC2A3* and *THBS*. However, this research showed several technical limitations and in particular the platform used and the different origin of the data.

In this study for the first time RNAseq and 450K methylation array were performed in a total of 39 primary Malignant Rhabdoid Tumours, investigating the epigenetic difference between these two malignancies. The current study was conducted to investigate gene signatures that drive Rhabdoid Tumours located in different parts of the body. The analysis was conducted in parallel with Medulloblastoma samples identified as a unique and homogenous group characterised by a common embryonal and neuronal gene signature and epigenetic events.

Our findings indicate that in Extra Cranial Rhabdoid and Atypical Teratoid Rhabdoid Tumours have common origins despite the location of the malignancies when compared with Medulloblastomas. Gene expression and methylation profiling reveals that Malignant Rhabdoid Tumours are a unique and coherent group from the other Medulloblastoma subgroups. In particular, genes associated with multiple processes such as cell proliferation, cell cycle, stem cell regulation and neurodevelopment emerged as differentially expressed and methylated compared with Medulloblastomas. This suggests that *SMARCB1* loss causes common deregulation of downstream biological mechanisms regardless of the tumour location.

The cohort was further interrogated to understand differences among the Rhabdoid Tumours. Gene expression and methylation profiling characterised two significant subgroups, in line with the previous findings. In particular, the analysis distinguished Extra Cranial Rhabdoid Tumours from Atypical Teratoid Rhabdoid Tumours. Extra Cranial Rhabdoid Tumours show significant enrichment and change in methylation profiling of genes associated muscle contraction and development, whilst in Atypical Teratoid Rhabdoid Tumours mainly genes involved in neurodevelopmental proliferation, such as *GFAP*, *B-FABP* and *SOX1*. Atypical Teratoid Rhabdoid Tumours also showed differentially expressed genes related to tumour progression, invasion, angiogenesis and hypermethylation of genes related to *Wnt* and *Notch* signalling pathways.

In conclusion the high-throughput profiling provided by this study showed that Rhabdoid Tumours are a unique biological identity compared with other CNS paediatric malignancies. Consistent with what has been previously published, the data sets also showed differences in gene expression and methylation status among Rhabdoid arising in different locations, indicating that the location of the tumours might be related to other genetic factors in conjunction with *SMARCB1* deletion.

### **8.2.1 Patterns of methylation and gene expression may have utility for disease sub-classification**

In this study we established that Rhabdoid Tumours are characterised by a distinct molecular signature, in comparison to other CNS paediatric tumours; nevertheless this high-throughput research characterised two different subtypes of malignancies mainly correlated with the primary location of the tumours, indicating the involvement of different molecular mechanisms. The two observed subgroups were validated in three independent trials cohorts, reported in chapter 6.

Rhabdoid Tumours are malignancies that are difficult to diagnose, since the mass is often located in the location typical of more common paediatric tumours. For instance,

Neuroblastoma, Nephroblastoma, Congenital Blastic Nephroma and Clear Cell Sarcoma of the kidney present foci similar to Extra Cranial Rhabdoid, while Atypical Teratoid Rhabdoid Tumours are often misdiagnosed as Medulloblastoma or CNS Primitive Neuroectodermal Tumour (PNET). Rhabdoid Tumours are now diagnosed by immunohistochemical staining, however absence or decrease of *SMARCB1* in immunohistochemistry reactivity is also observed in Renal Medullary Carcinoma, Epithelioid sarcoma, poorly differentiated Synovial Sarcoma, Ewing sarcoma Schwannomatosis (Rousseau, Noguchi et al. 2011), Chordomas (Mobley, McKenney et al. 2010) and Choroid Plexus Carcinoma.

In this study I have indicated the utility of DNA methylation and next generation sequencing not just for disease sub-classification, but also for diagnostics. These findings provide insights into the molecular pathways involved in Rhabdoid tumourigenesis and will contribute to a better understanding, and treatment of these malignancies. Also the use of molecular profiling in clinics will improve diagnostic and in particular to avoid misinterpreting SMARCB1 immuno-reactivity or tumour location.

### 8.3 Epigenetic effect of *SMARCB1* deletion in Rhabdoid Tumours

Malignant Rhabdoid Tumours are characterised by a biallelic inactivation at chromosome 22 of *SMARCB1* (Jackson, Sievert et al. 2009). *SMARCB1* is indicated as potent tumour suppressor and its implication in tumourigenesis has been previously demonstrated both *in vivo* and *in vitro*. In particular, re-expression of *SMARCB1* in Rhabdoid cell lines determines deregulation of genes related with cell cycle and cell proliferation (Medjkane, Novikov et al. 2004, Morozov, Lee et al. 2007).

To further investigate the role of *SMARCB1* in tumourigenesis, for the first time RNA-seq and 450K were performed in Rhabdoid cell lines in which *SMARCB1* was re-expressed and also treated with 5-aza-2'-deoxycytidine. The research aimed to understand the effect of *SMARCB1* loss in the epigenetic mechanism and characterise genes and pathway directly dependent on *SMARCB1* that could be targeted in Malignant Rhabdoid Tumours

The investigation in *SMARCB1* re-expressing cells indicated that *SMARCB1* causes mainly hypomethylation and upregulation of a wide number of genes involved in neurodevelopment and promoting metastasis and angiogenesis. Interestingly, we observed that *SMARCB1* re-expression causes gene deregulation by provoking alternative splicing. RNAseq analysis also indicates that *SMARCB1* deregulates a significant number of genes and isoforms, whose role needs to be defined.

Treatment of Rhabdoid cell lines with 5-aza-2'-deoxycytidine showed a general upregulation of genes involved in chromatin stability and in modulation of Polycomb target genes. Combining data from primary and functional models (*SMARCB1* re-expression/5-aza-CdR treatment) highlighted 13 genes which were appropriately deregulated in all conditions, suggesting a *SMARCB1* demethylating effect. In fact, cross-referencing across platform and data sets from primary samples and *SMARCB1* re-expressing cell lines suggests that *SMARCB1* may cause gene deregulation by directly modulating methylation modifications. In particular, direct upregulation caused by hypomethylation was observed in *GLI1*, *GLI3*, *JARID2*, *NTN4*, *GALNT2*, *BHLHE4* and *ABR*.



In conclusion, this study investigated the epigenetic mechanism of *SMARCB1* and characterised downstream effectors of *SMARCB1* that can be targeted for arresting the growth of Rhabdoid cells. In particular, we show that *SMARCB1* is actively involved in the methylation machinery and its deletion causes deregulation of expression of a variety of gene and also activates alternative splicing.

#### **8.4 Integrative analysis of model and primary data indicate activation of a variety of pathways dependent on *SMARCB1* deletion**

Previous profiling of primary Rhabdoid Tumours and *SMARCB1* re-expressing Rhabdoid cell lines indicated that multiple pathways are deregulated following *SMARCB1* deletion. However, only a few genes such as *Cyclin D1* were suggested as possible drug target in Rhabdoid Tumours in children.

In this study for the first time, cross-referencing between data sets was strategically used to characterise *SMARCB1*-dependent Rhabdoid tumour gene signatures that may be used as therapeutic targets. In particular we identified 13 genes which expression was re-established after both *SMARCB1* re-expression and 5-aza-2'-deoxycytidine treatment and correlated with non Rhabdoid Tumours. All 13 genes are cancer-related and for most of them the tumourigenic mechanism is still not completely understood. Interestingly Dnm3os (gene transcribed into a non-coding RNA that contains three micro RNAs (miRNAs), miR-199a, miR-199a, and miR-214) was identified as part of this set, suggesting that *SMARCB1* might also recruit SWI/SNF complexes at miRNA promoters.

A comparison of gene signatures between functional models and primary samples also provided a rational way to analyse pathways involved in Rhabdoid tumourigenesis. The investigation identified a variety of canonical pathways and biological functions that were deregulated in Rhabdoid primary tumours and re-established following *SMARCB1* induction and 5-aza-2'-deoxycytidine treatment. Specifically the analysis distinguished pathways and upstream regulators mainly involved in cell cycle progression, recruitment of SWI/SNF complexes, Interferon Gamma, RhoA signalling, TP53 signalling, etc. Finally, molecular activity prediction analysis based on ingenuity annotation identified TGFbeta as a likely significant upstream regulator of the tumourigenic pathway in Rhabdoid Tumours.

In conclusion, this study has indicated the importance of cross-referencing model data and primary profiling to identify genes and pathways that drive tumourigenesis in Rhabdoid Tumours. A subset of genes were identified as *SMARCB1*/methylation dependent genes and also correlated with a Rhabdoid phenotype. Moreover pathway analysis indicates deregulation of several pathways and upstream regulator involved in Rhabdoid tumourigenesis that will be further investigated to understand their potential as candidates for developing targeted therapy.

## 8.5 Limitations

This study investigated the role of *SMARCB1* both in cells and primary tumours. In particular *SMARCB1* was re-expressed in Rhabdoid cell lines using a lentivirus approach. Re-expression or overexpression of genes is a technique used extensively in genetic studies. However, the mRNA level of the re-expressed genes is often higher than the biological samples and level of reexpression inconsistent amongst replicates. Additionally, in this study the lentiviral approaches used CMV as constitutive promoter that induces continual induction of the genes of interest, contrary from what normally happens physiologically. However, the cell line STA-WT1 which maintains *SMARCB1* protein (even if in a lower levels) and the effect on gene expression caused by 5-aza-2'-deoxycytidine treatment suggested that the effect of *SMARCB1* re-expression was not simply due to expression at a non-physiological level. Nevertheless, inducible promoters would improve the experimental condition, since the intensity of the gene of interested could be modulated.

qPCR and Western Blot analysis show an increase of the *SMARCB1* mRNA levels and protein levels following transduction, however this study is limited in its ability to provide evidence of incorporation of SMARCB1 protein into the SWI/SNF complex. Despite ChIP results showing a change in recruitment of SMARCA4 at the gene promoter following *SMARCB1* reexpression, this does not fully confirm the involvement of *SMARCB1* in the functional and physiological activity of the SWI/SNF complexes. Comparison of SMARCB1, SMARCA4 and other constitutive subunits binding measured by ChIP-seq will give a better understanding of the dynamic recruitment of the SWI/SNF complex in Malignant Rhabdoid Tumours.

Introduction of both *SMARCA4* D/N and WT in Rhabdoid cell lines causes activation of apoptosis and upregulation of *CDKN2A* and *CDKN1A* in A204 and G401 cell lines. However this effect typically had to be measured at an early time point and so the effect could be caused by Puromycin selective treatment. This also is balanced by the fact that cells in which *SMARCA4* WT was induced did not show changes in morphology nor arrest of cell growth in longer-term culture. Unfortunately time did not allow for further

analysis or developing an inducible system. The molecular effect of mutant *SMARCA4* in Rhabdoid cell lines should be further investigated.

The profiling study on primary tumours aimed to understand the mechanism of tumourigenesis and characterise the differences between Rhabdoid Tumours based on the location of the malignancies. Despite the fact that all Malignant Rhabdoid Tumours cluster in a different metagene when compared with Medulloblastoma, the data suggested that the cohort is not simply composed of two classes purely divided by tumour location. In fact, projection of the metagene that characterised Extra Cranial Rhabdoid tumours and Atypical Teratoid Rhabdoid Tumours onto other platforms further supports the non-exclusivity of the two groups i.e. some Atypical Teratoid Rhabdoid resemble Extra Cranial Rhabdoid tumours molecularly and *vice versa*. The presence of a further third molecular subgroup should also not be ruled out: some Extra Cranial Rhabdoid tumours and Atypical Teratoid Rhabdoid Tumours appear to present mixed features showing gene signatures of both the bigger subgroups. In our cohort of 23 primary samples just 2 presented this phenotype. Studies of a larger cohort may provide a more accurate ability to detect this possible third group and finally to characterise it.

## **8.6 Future work**

### **8.6.1 Implication of gene deregulation in Rhabdoid Tumourigenesis**

The present study aimed to collect a catalogue of data to inform as to critical genes and pathways involved in Rhabdoid tumourigenesis. This work was able to characterise the epigenetic modification caused by *SMARCB1* loss and extract gene signatures that might explain the mechanism behind Rhabdoid malignancies. Despite the validation of some of the findings, further investigation will be required to establish the role of the deregulated genes extracted from this comparative analysis. High-throughput genetic screening perhaps with a CRISPR library may also be useful to identify genes that are involved in Rhabdoid tumourigenesis.

### **8.6.2 Characterisation of pathways**

Comparative analysis between platform and data sets provided a method to identified possible pathways involved in cancer development in Rhabdoid Tumours. However deregulation of gene expression does not always causes changes at the protein level, and it also has been observed that upregulation of mRNA may also cause downregulation of the protein effector (Cham, Xu et al. 2003). Further investigation of the changes in protein levels as results of *SMARCB1* re-expression in Rhabdoid cell lines will be performed by proteomic study. In particular, stable isotope labelling proteomics in Rhabdoid cell lines re-expressing *SMARCB1* will be used to separate, identify and analyse those proteins affected by *SMARCB1*. This approach will give us the possibility to compare protein levels and gene expression and finally determine the tumourigenic pathways involved in Rhabdoid Tumours.

### **8.6.3 Characterisation of *SMARCA4* D/N effect in gene expression and methylation status**

A preliminary study of *SMARCA4* role was attempted. Here we showed that *SMARCA4* D/N induces apoptosis and a rapid cell death. However some increase in apoptosis was also observed when the wild type form was induced. It is still not clear if the effect was caused by the early time point of the experiment and no analysis of the molecular biology (RNA-seq) has yet been performed. An inducible system will provide a better model to study, since no process of selection is required. Moreover, RNA-seq and 450K methylation of Rhabdoid cells with induction of *SMARCA4* D/N will be performed and compared to the experimental data set of this study.

### **8.6.4 Single cell sequencing**

Malignant Rhabdoid Tumours and in particular Atypical Teratoid Rhabdoid Tumours present a neuroepithelial, peripheral epithelial, and mesenchymal element in addition to Rhabdoid features. Moreover, in this study it can also be indicated that a small subset of primary tumours showed a mixed genotype that may reassemble both the two major subgroup we have characterised. Moreover, *SMARCB1* is not necessarily completely deleted in all Rhabdoid Tumours, and low levels of mutated protein may still be detected. Single cell sequencing will provide a better understanding of the many biological processes that occur in individual cells and better characterise the Rhabdoid Tumours we observe. Moreover, this will provide an analysis of the functions of individual cells in the context of its microenvironment, which might explain the development of these tumours in such different locations of the body.

## 8.7 Final Conclusion

This project aimed to characterise the effects of *SMARCB1* in Rhabdoid cell lines and identified genes that might be therapeutically targeted. High throughput sequencing and methylation array were used to outline a better picture of the role of *SMARCB1* in gene expression modulation. Moreover cross-referencing between platforms and primary and experimental data was applied to find gene signatures that resemble a non Rhabdoid phenotype and understand the effect on methylation of *SMARCB1* re-expression in Rhabdoid Tumours.

This investigation has made a substantial contribution to knowledge in the field by defining the gene expression and methylation characteristics of primary Malignant Rhabdoid Tumours and defined the extent to which their gene expression and methylation profile is dependent upon *SMARCB1* loss. In this study I have indicated that whilst Malignant Rhabdoid Tumours form a single biological entity they can be further sub-grouped into at least two groups. Also here I show the direct effect on methylation status caused by *SMARCB1* re-expression and the consequential change in expression, in a similar manner we observed following 5-aza-2'-deoxycytidine. I have catalogued a series of *SMARCB1*-dependent genes as well as pathways, which may now be further explored as potential novel therapeutic targets.

Furthermore, we have indicated that next generation profiling of primary tumours are useful in patient stratification, and in particular in diagnosis of all paediatric tumours that present *SMARCB1* deletion and are phenotypically similar to Malignant Rhabdoid Tumours.



## **Chapter 9      References**



Moving AHEAD with an international human epigenome project. (2008). *Nature*, 454(7205), 711-715.

Ae, K., Kobayashi, N., Sakuma, R., Ogata, T., Kuroda, H., Kawaguchi, N., Kitamura, Y. (2002). Chromatin remodelling factor encoded by *ini1* induces G1 arrest and apoptosis in *ini1*-deficient cells. *Oncogene*, 21(20), 3112-3120. doi: 10.1038/sj.onc.1205414

Ahringer, J. (2000). NuRD and SIN3 histone deacetylase complexes in development. *Trends Genet*, 16(8), 351-356.

Akhbari, P., Whitehouse, A., & Boyne, J. R. (2014). Long non-coding RNAs drive metastatic progression in melanoma (Review). *Int J Oncol*, 45(6), 2181-2186. doi: 10.3892/ijo.2014.2691

Akhurst, R. J., & Derynck, R. (2001). TGF-beta signaling in cancer--a double-edged sword. *Trends Cell Biol*, 11(11), S44-51.

Alarcon-Vargas, D., Zhang, Z., Agarwal, B., Challagulla, K., Mani, S., & Kalpana, G. V. (2006). Targeting cyclin D1, a downstream effector of INI1/hSNF5, in rhabdoid tumors. *Oncogene*, 25(5), 722-734. doi: 10.1038/sj.onc.1209112

Albanese, P., Belin, M. F., & Delattre, O. (2006). The tumour suppressor hSNF5/INI1 controls the differentiation potential of malignant rhabdoid cells. *Eur J Cancer*, 42(14), 2326-2334. doi: 10.1016/j.ejca.2006.03.028

Albini, S., & Puri, P. L. (2010). SWI/SNF complexes, chromatin remodeling and skeletal myogenesis: it's time to exchange! *Exp Cell Res*, 316(18), 3073-3080. doi: 10.1016/j.yexcr.2010.05.023

Alexaki, V. I., Javelaud, D., Van Kempen, L. C., Mohammad, K. S., Dennler, S., Luciani, F., Mauviel, A. (2010). GLI2-mediated melanoma invasion and metastasis. *J Natl Cancer Inst*, 102(15), 1148-1159. doi: 10.1093/jnci/djq257

- Alholle, A., Brini, A. T., Gharanei, S., Vaiyapuri, S., Arrigoni, E., Dallol, A., Latif, F. (2013). Functional epigenetic approach identifies frequently methylated genes in Ewing sarcoma. *Epigenetics*, 8(11), 1198-1204. doi: 10.4161/epi.26266
- Ali, I., Haque, A., Wani, W. A., Saleem, K., & Al Za'abi, M. (2013). Analyses of anticancer drugs by capillary electrophoresis: A review. *Biomedical Chromatography*, 27(10), 1296-1311.
- Alimova, I., Birks, D. K., Harris, P. S., Knipstein, J. A., Venkataraman, S., Marquez, V. E., Vibhakar, R. (2013). Inhibition of EZH2 suppresses self-renewal and induces radiation sensitivity in atypical rhabdoid teratoid tumor cells. *Neuro Oncol*, 15(2), 149-160. doi: 10.1093/neuonc/nos285
- Allen, J. C., Judkins, A. R., Rosenblum, M. K., & Biegel, J. A. (2006). Atypical teratoid/rhabdoid tumor evolving from an optic pathway ganglioglioma: case study. *Neuro Oncol*, 8(1), 79-82. doi: 10.1215/s1522851705000347
- Al-Samadi, A., Kouri, V.-P., Salem, A., Ainola, M., Kaivosoja, E., Barreto, G., Häyrinen-Immonen, R. (2014). IL-17C and its receptor IL-17RA/IL-17RE identify human oral epithelial cell as an inflammatory cell in recurrent aphthous ulcer. *Journal of Oral Pathology & Medicine*, 43(2), 117-124. doi: 10.1111/jop.12095
- Amankwah, E. K., Thompson, R. C., Nabors, L. B., Olson, J. J., Browning, J. E., Madden, M. H., & Egan, K. M. (2013). SWI/SNF gene variants and glioma risk and outcome. *Cancer Epidemiol*, 37(2), 162-165. doi: 10.1016/j.canep.2012.12.001
- Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol*, 11(10), R106. doi: 10.1186/gb-2010-11-10-r106
- Anders, S., Pyl, P. T., & Huber, W. (2014). HTSeq - A Python framework to work with high-throughput sequencing data.
- Anders, S., Reyes, A., & Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *Genome Res*, 22(10), 2008-2017. doi: 10.1101/gr.133744.111

- Anderton, J. A., Lindsey, J. C., Lusher, M. E., Gilbertson, R. J., Bailey, S., Ellison, D. W., & Clifford, S. C. (2008). Global analysis of the medulloblastoma epigenome identifies disease-subgroup-specific inactivation of COL1A2. *Neuro Oncol*, 10(6), 981-994. doi: 10.1215/15228517-2008-048
- Aoyagi, S., & Hayes, J. J. (2002). hSWI/SNF-catalyzed nucleosome sliding does not occur solely via a twist-diffusion mechanism. *Mol Cell Biol*, 22(21), 7484-7490.
- Appolloni, I., Calzolari, F., Barilari, M., Terrile, M., Daga, A., & Malatesta, P. (2012). Antagonistic modulation of gliomagenesis by Pax6 and Olig2 in PDGF-induced oligodendroglioma. *Int J Cancer*, 131(7), E1078-1087. doi: 10.1002/ijc.27606
- Archacki, R., Buszewicz, D., Sarnowski, T. J., Sarnowska, E., Rolicka, A. T., Tohge, T., Jerzmanowski, A. (2013). BRAHMA ATPase of the SWI/SNF chromatin remodeling complex acts as a positive regulator of gibberellin-mediated responses in arabidopsis. *PLoS One*, 8(3), e58588. doi: 10.1371/journal.pone.0058588
- Armeanu-Ebinger, S., Herrmann, D., Bonin, M., Leuschner, I., Warmann, S. W., Fuchs, J., & Seitz, G. (2012). Differential expression of miRNAs in rhabdomyosarcoma and malignant rhabdoid tumor. *Exp Cell Res*, 318(20), 2567-2577. doi: <http://dx.doi.org/10.1016/j.yexcr.2012.07.015>
- Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, 30(10), 1363-1369. doi: 10.1093/bioinformatics/btu049
- Ashoor, H., Hérault, A., Kamoun, A., Radvanyi, F., Bajic, V. B., Barillot, E., & Boeva, V. (2013). HMCan: a method for detecting chromatin modifications in cancer samples using ChIP-seq data. *Bioinformatics*, 29(23), 2979-2986. doi: 10.1093/bioinformatics/btt524

- Atsuta, Y., Aoki, N., Sato, K., Oikawa, K., Nochi, H., Miyokawa, N., Katagiri, M. (2002). Identification of metalloproteinase-1 as a member of a tumor associated antigen in patients with breast cancer. *Cancer Lett*, 182(1), 101-107.
- Bae, S., Bessho, Y., Hojo, M., & Kageyama, R. (2000). The bHLH gene Hes6, an inhibitor of Hes1, promotes neuronal differentiation. *Development*, 127(13), 2933-2943.
- Bagchi, A., Papazoglu, C., Wu, Y., Capurso, D., Brodt, M., Francis, D., Mills, A. A. (2007). CHD5 Is a Tumor Suppressor at Human 1p36. *Cell*, 128(3), 459-475.
- Bai, X. Y., Lin, J. Y., Zhang, X. C., Xie, Z., Yan, H. H., Chen, Z. H., Wu, Y. L. (2013). High expression of truncated GLI3 is associated with poor overall survival in patients with non-small cell lung cancer. *Cancer Biomark*, 13(1), 37-47. doi: 10.3233/cbm-130312
- Banderli, U., Jayanthan, A., Hoeksema, K. A., Narendran, A., & Giles, W. R. (2012). Ion channels in pediatric CNS Atypical Teratoid/Rhabdoid Tumor (AT/RT) cells: potential targets for novel therapeutic agents. *J Neurooncol*, 107(1), 111-119. doi: 10.1007/s11060-011-0735-x
- Banine, F., Bartlett, C., Gunawardena, R., Muchardt, C., Yaniv, M., Knudsen, E. S., Sherman, L. S. (2005). SWI/SNF chromatin-remodeling factors induce changes in DNA methylation to promote transcriptional activation. *Cancer Res*, 65(9), 3542-3547. doi: 10.1158/0008-5472.can-04-3554
- Bansal, K. K., & Goel, D. (2007). Atypical teratoid/rhabdoid tumors of central nervous system. *J Paediatr Neurosc*, 2(1), 309-313. doi: 10.1126/science.1141319
- Barker, N., Hurlstone, A., Musisi, H., Miles, A., Bienz, M., & Clevers, H. (2001). The chromatin remodelling factor Brg-1 interacts with beta-catenin to promote target gene activation. *Embo j*, 20(17), 4935-4943. doi: 10.1093/emboj/20.17.4935

Bartelheim, K., Sumerauer, D., Behrends, U., Kodetova, D., Kucera, F., Leuschner, I., Fruhwald, M. C. (2014). Clinical and genetic features of rhabdoid tumors of the heart registered with the European Rhabdoid Registry (EU-RHAB). *Cancer Genet.* doi: 10.1016/j.cancergen.2014.04.005

Baylin, S. B., & Herman, J. G. (2000). DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet*, 16(4), 168-174.

Baylin, S. B., & Jones, P. A. (2011). A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer*, 11(10), 726-734. doi: 10.1038/nrc3130

Becker, T. M., Haferkamp, S., Dijkstra, M. K., Scurr, L. L., Frausto, M., Diefenbach, E., Rizos, H. (2009). The chromatin remodelling factor BRG1 is a novel binding partner of the tumor suppressor p16INK4a. *Mol Cancer*, 8, 4. doi: 10.1186/1476-4598-8-4

Bellon, N., Fraitag, S., Miquel, C., Salomon, L. J., Bourdeaut, F., Bodemer, C., Hadj-Rabia, S. (2014). Cutaneous location of atypical teratoid/rhabdoid tumour. *Acta Derm Venereol*, 94(4), 454-456. doi: 10.2340/00015555-1716

Benesch, M., Bartelheim, K., Fleischhack, G., Gruhn, B., Schlegel, P. G., Witt, O., Fruhwald, M. C. (2014). High-dose chemotherapy (HDCT) with auto-SCT in children with atypical teratoid/rhabdoid tumors (AT/RT): a report from the European Rhabdoid Registry (EU-RHAB). *Bone Marrow Transplant*, 49(3), 370-375. doi: 10.1038/bmt.2013.208

Bestor, T. H. (1990). DNA methylation: evolution of a bacterial immune function into a regulator of gene expression and genome structure in higher eukaryotes. *Philos Trans R Soc Lond B Biol Sci*, 326(1235), 179-187.

Betz, B. L., Strobeck, M. W., Reisman, D. N., Knudsen, E. S., & Weissman, B. E. (2002). Re-expression of hSNF5/INI1/BAF47 in pediatric tumor cells leads to G1 arrest associated with induction of p16ink4a and activation of RB. *Oncogene*, 21(34), 5193-5203. doi: 10.1038/sj.onc.1205706

- Bevilacqua, A., Willis, M. S., & Bultman, S. J. (2014). SWI/SNF chromatin-remodeling complexes in cardiovascular development and disease. *Cardiovasc Pathol*, 23(2), 85-91. doi: 10.1016/j.carpath.2013.09.003
- Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J. M., Shen, R. (2011). High density DNA methylation array with single CpG site resolution. *Genomics*, 98(4), 288-295. doi: <http://dx.doi.org/10.1016/j.ygeno.2011.07.007>
- Biegel, J. A. (1997). Genetics of pediatric central nervous system tumors. *J Pediatr Hematol Oncol*, 19(6), 492-501.
- Biegel, J. A. (1999). Cytogenetics and molecular genetics of childhood brain tumors. *Neuro Oncol*, 1(2), 139-151.
- Biegel, J. A. (2006). Molecular genetics of atypical teratoid/rhabdoid tumor. *Neurosurg Focus*, 20(1), E11.
- Biegel, J. A., Busse, T. M., & Weissman, B. E. (2014). SWI/SNF chromatin remodeling complexes and cancer. *Am J Med Genet C Semin Med Genet*, 166(3), 350-366. doi: 10.1002/ajmg.c.31410
- Biegel, J. A., Fogelgren, B., Wainwright, L. M., Zhou, J. Y., Bevan, H., & Rorke, L. B. (2000). Germline INI1 mutation in a patient with a central nervous system atypical teratoid tumor and renal rhabdoid tumor. *Genes Chromosomes Cancer*, 28(1), 31-37.
- Biegel, J. A., Kalpana, G., Knudsen, E. S., Packer, R. J., Roberts, C. W., Thiele, C. J., Smith, M. (2002). The role of INI1 and the SWI/SNF complex in the development of rhabdoid tumors: meeting summary from the workshop on childhood atypical teratoid/rhabdoid tumors. *Cancer Res*, 62(1), 323-328.
- Biegel, J. A., Tan, L., Zhang, F., Wainwright, L., Russo, P., & Rorke, L. B. (2002). Alterations of the hSNF5/INI1 gene in central nervous system atypical teratoid/rhabdoid tumors and renal and extrarenal rhabdoid tumors. *Clin Cancer Res*, 8(11), 3461-3467.



Biegel, J. A., Zhou, J. Y., Rorke, L. B., Stenstrom, C., Wainwright, L. M., & Fogelgren, B. (1999). Germ-line and acquired mutations of INI1 in atypical teratoid and rhabdoid tumors. *Cancer Res*, 59(1), 74-79.

Bikowska, B., Grajkowska, W., & Jozwiak, J. (2011). Atypical teratoid/rhabdoid tumor: short clinical description and insight into possible mechanism of the disease. *Eur J Neurol*, 18(6), 813-818. doi: 10.1111/j.1468-1331.2010.03277.x

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev*, 16(1), 6-21. doi: 10.1101/gad.947102

Birks, D. K., Donson, A. M., Patel, P. R., Dunham, C., Muscat, A., Algar, E. M., Foreman, N. K. (2011). High expression of BMP pathway genes distinguishes a subset of atypical teratoid/rhabdoid tumors associated with shorter survival. *Neuro Oncol*, 13(12), 1296-1307. doi: 10.1093/neuonc/nor140

Birks, D. K., Donson, A. M., Patel, P. R., Sufit, A., Algar, E. M., Dunham, C., Foreman, N. K. (2013). Pediatric rhabdoid tumors of kidney and brain show many differences in gene expression but share dysregulation of cell cycle and epigenetic effector genes. *Pediatr Blood Cancer*, 60(7), 1095-1102. doi: 10.1002/pbc.24481

Birks, D. K., Kleinschmidt-DeMasters, B. K., Donson, A. M., Barton, V. N., McNatt, S. A., Foreman, N. K., & Handler, M. H. (2010). Claudin 6 is a positive marker for atypical teratoid/rhabdoid tumors. *Brain Pathol*, 20(1), 140-150. doi: 10.1111/j.1750-3639.2008.00255.x

Birks, D. K., Kleinschmidt-DeMasters, B. K., Donson, A. M., Barton, V. N., McNatt, S. A., Foreman, N. K., & Handler, M. H. (2010). Claudin 6 Is a Positive Marker for Atypical Teratoid/Rhabdoid Tumors. *Brain Pathology*, 20(1), 140-150. doi: 10.1111/j.1750-3639.2008.00255.x

Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., Margulies, E. H., de Jong, P. J. (2007). Identification and analysis of functional elements in 1% of

the human genome by the ENCODE pilot project. *Nature*, 447(7146), 799-816. doi: 10.1038/nature05874

Bochar, D. A., Wang, L., Beniya, H., Kinev, A., Xue, Y., Lane, W. S., Shiekhattar, R. (2000). BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. *Cell*, 102(2), 257-265.

Bogdanovic, O., & Veenstra, G. J. (2009). DNA methylation and methyl-CpG binding proteins: developmental requirements and function. *Chromosoma*, 118(5), 549-565. doi: 10.1007/s00412-009-0221-9

Bonazzi, V., Medjkane, S., Quignon, F., & Delattre, O. (2005). Complementation analyses suggest species-specific functions of the SNF5 homology domain. *Biochem Biophys Res Commun*, 336(2), 634-638. doi: 10.1016/j.bbrc.2005.08.133

Booth, M. J., Branco, M. R., Ficz, G., Oxley, D., Krueger, F., Reik, W., & Balasubramanian, S. (2012). Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science*, 336(6083), 934-937. doi: 10.1126/science.1220671

Bosse, K. R., Shukla, A. R., Pawel, B., Chikwava, K. R., Santi, M., Tooke, L., Bagatell, R. (2014). Malignant rhabdoid tumor of the bladder and ganglioglioma in a 14 year-old male with a germline 22q11.2 deletion. *Cancer Genet*. doi: 10.1016/j.cancergen.2014.05.007

Bourdeaut, F., & Bieche, I. (2012). [Chromatin remodeling defects and cancer: the SWI/SNF example]. *Bull Cancer*, 99(12), 1133-1140. doi: 10.1684/bdc.2012.1664

Bourdeaut, F., Dufour, C., & Delattre, O. (2010). [Rhabdoid tumours: hSNF/INI1 deficient cancers of early childhood with aggressive behaviour]. *Bull Cancer*, 97(1), 37-45. doi: 10.1684/bdc.2009.1024

- Bourdeaut, F., Freneaux, P., Thuille, B., Bergeron, C., Laurence, V., Brugieres, L., Orbach, D. (2008). Extra-renal non-cerebral rhabdoid tumours. *Pediatr Blood Cancer*, 51(3), 363-368. doi: 10.1002/pbc.21632
- Bourdeaut, F., Freneaux, P., Thuille, B., Lellouch-Tubiana, A., Nicolas, A., Couturier, J., Delattre, O. (2007). hSNF5/INI1-deficient tumours and rhabdoid tumours are convergent but not fully overlapping entities. *J Pathol*, 211(3), 323-330. doi: 10.1002/path.2103
- Bourdeaut, F., Lequin, D., Brugieres, L., Reynaud, S., Dufour, C., Doz, F., Delattre, O. (2011). Frequent hSNF5/INI1 germline mutations in patients with rhabdoid tumor. *Clin Cancer Res*, 17(1), 31-38. doi: 10.1158/1078-0432.ccr-10-1795
- Bowen, N. J., Fujita, N., Kajita, M., & Wade, P. A. (2004). Mi-2/NuRD: Multiple complexes for many purposes. *Biochimica et Biophysica Acta - Gene Structure and Expression*, 1677(1-3), 52-57.
- Boyer, L. A., Latek, R. R., & Peterson, C. L. (2004). The SANT domain: A unique histone-tail-binding module? *Nature Reviews Molecular Cell Biology*, 5(2), 158-163.
- Brennan, B., Stiller, C., & Bourdeaut, F. (2013). Extracranial rhabdoid tumours: what we have learned so far and future directions. *Lancet Oncol*, 14(8), e329-336. doi: 10.1016/s1470-2045(13)70088-3
- Brodská, B., Holoubek, A., Otevřelová, P., & Kuželová, K. (2013). Combined treatment with low concentrations of decitabine and saha causes cell death in leukemic cell lines but not in normal peripheral blood lymphocytes. *BioMed Research International*, 2013.
- Brown, T. (2002). *Genomes*. 2nd edition Wiley-Liss (Ed.) Chapter 3, Transcriptomes and Proteomes Retrieved from <http://dx.doi.org/10.1038/nrg2484>
- Bruder, C. E., Dumanski, J. P., & Kedra, D. (1999). The mouse ortholog of the human SMARCB1 gene encodes two splice forms. *Biochem Biophys Res Commun*, 257(3), 886-890. doi: 10.1006/bbrc.1999.0563

- Bruggers, C. S., Bleyl, S. B., Pysher, T., Barnette, P., Afify, Z., Walker, M., & Biegel, J. A. (2011). Clinicopathologic comparison of familial versus sporadic atypical teratoid/rhabdoid tumors (AT/RT) of the central nervous system. *Pediatr Blood Cancer*, 56(7), 1026-1031. doi: 10.1002/pbc.22757
- Brunet, J. P., Tamayo, P., Golub, T. R., & Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A*, 101(12), 4164-4169. doi: 10.1073/pnas.0308531101
- Brunet, J.-P., Tamayo, P., Golub, T. R., & Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences*, 101(12), 4164-4169. doi: 10.1073/pnas.0308531101
- Buanne, P., Corrente, G., Micheli, L., Palena, A., Lavia, P., Spadafora, C., Tirone, F. (2000). Cloning of PC3B, a Novel Member of the PC3/BTG/TOB Family of Growth Inhibitory Genes, Highly Expressed in the Olfactory Epithelium. *Genomics*, 68(3), 253-263. doi: <http://dx.doi.org/10.1006/geno.2000.6288>
- Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A., Magnuson, T. (2000). A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Mol Cell*, 6(6), 1287-1295.
- Burger, P. C., Yu, I. T., Tihan, T., Friedman, H. S., Strother, D. R., Kepner, J. L., Perlman, E. J. (1998). Atypical teratoid/rhabdoid tumor of the central nervous system: a highly malignant tumor of infancy and childhood frequently mistaken for medulloblastoma: a Pediatric Oncology Group study. *Am J Surg Pathol*, 22(9), 1083-1092.
- Buscariollo, D. L., Park, H. S., Roberts, K. B., & Yu, J. B. (2012). Survival outcomes in atypical teratoid rhabdoid tumor for patients undergoing radiotherapy in a Surveillance, Epidemiology, and End Results analysis. *Cancer*, 118(17), 4212-4219. doi: 10.1002/cncr.27373

- Cairns, J. M., Dunning, M. J., Ritchie, M. E., Russell, R., & Lynch, A. G. (2008). BASH: a tool for managing BeadArray spatial artefacts. *Bioinformatics*, 24(24), 2921-2922. doi: 10.1093/bioinformatics/btn557
- Campbell, M. J., & Turner, B. M. (2013). Altered histone modifications in cancer. *Adv Exp Med Biol*, 754, 81-107. doi: 10.1007/978-1-4419-9967-2\_4
- Caracciolo, V., D'Agostino, L., Draberova, E., Sladkova, V., Crozier-Fitzgerald, C., Agamanolis, D. P., Katsetos, C. D. (2010). Differential expression and cellular distribution of gamma-tubulin and betaIII-tubulin in medulloblastomas and human medulloblastoma cell lines. *J Cell Physiol*, 223(2), 519-529. doi: 10.1002/jcp.22077
- Caramel, J., Medjkane, S., Quignon, F., & Delattre, O. (2008). The requirement for SNF5/INI1 in adipocyte differentiation highlights new features of malignant rhabdoid tumors. *Oncogene*, 27(14), 2035-2044. doi: 10.1038/sj.onc.1210847
- Caramel, J., Quignon, F., & Delattre, O. (2008). RhoA-dependent regulation of cell migration by the tumor suppressor hSNF5/INI1. *Cancer Res*, 68(15), 6154-6161. doi: 10.1158/0008-5472.can-08-0115
- Castelo-Branco, P., Choufani, S., Mack, S., Gallagher, D., Zhang, C., Lipman, T., Tabori, U. (2013). Methylation of the TERT promoter and risk stratification of childhood brain tumours: an integrative genomic and molecular study. *Lancet Oncol*, 14(6), 534-542. doi: [http://dx.doi.org/10.1016/S1470-2045\(13\)70110-4](http://dx.doi.org/10.1016/S1470-2045(13)70110-4)
- Chai, J., Charboneau, A. L., Betz, B. L., & Weissman, B. E. (2005). Loss of the hSNF5 gene concomitantly inactivates p21CIP/WAF1 and p16INK4a activity associated with replicative senescence in A204 rhabdoid tumor cells. *Cancer Res*, 65(22), 10192-10198. doi: 10.1158/0008-5472.can-05-1896
- Cham, C. M., Xu, H., O'Keefe, J. P., Rivas, F. V., Zagouras, P., & Gajewski, T. F. (2003). Gene array and protein expression profiles suggest post-transcriptional regulation during CD8+ T cell differentiation. *J Biol Chem*, 278(19), 17044-17052. doi: 10.1074/jbc.M212741200

- Chan, K. H., Mohammed Haspani, M. S., Tan, Y. C., & Kassim, F. (2011). A case report of atypical teratoid/rhabdoid tumour in a 9-year-old girl. *Malays J Med Sci*, 18(3), 82-86.
- Chan, W. M., Rahman, M. M., & McFadden, G. (2013). Oncolytic myxoma virus: the path to clinic. *Vaccine*, 31(39), 4252-4258. doi: 10.1016/j.vaccine.2013.05.056
- Chandler, S. P., Guschin, D., Landsberger, N., & Wolffe, A. P. (1999). The Methyl-CpG Binding Transcriptional Repressor MeCP2 Stably Associates with Nucleosomal DNA. *Biochemistry*, 38(22), 7008-7018. doi: 10.1021/bi990224y
- Chandrakesan, P., Weygant, N., May, R., Qu, D., Chinthalapally, H. R., Sureban, S. M., Houchen, C. W. (2014). DCLK1 facilitates intestinal tumor growth via enhancing pluripotency and epithelial mesenchymal transition. *Oncotarget*, 5(19), 9269-9280.
- Chen, M. L., McComb, J. G., & Krieger, M. D. (2005). Atypical teratoid/rhabdoid tumors of the central nervous system: management and outcomes. *Neurosurg Focus*, 18(6a), E8.
- Cheng, S. W., Davies, K. P., Yung, E., Beltran, R. J., Yu, J., & Kalpana, G. V. (1999). c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function. *Nat Genet*, 22(1), 102-105. doi: 10.1038/8811
- Cheng, Y. C., Lirng, J. F., Chang, F. C., Guo, W. Y., Teng, M. M., Chang, C. Y., Ho, D. M. (2005). Neuroradiological findings in atypical teratoid/rhabdoid tumor of the central nervous system. *Acta Radiol*, 46(1), 89-96.
- Cheung, W. K., Zhao, M., Liu, Z., Stevens, L. E., Cao, P. D., Fang, J. E., Nguyen, D. X. (2013). Control of alveolar differentiation by the lineage transcription factors GATA6 and HOPX inhibits lung adenocarcinoma metastasis. *Cancer Cell*, 23(6), 725-738. doi: 10.1016/j.ccr.2013.04.009

- Chi, S. N., Zimmerman, M. A., Yao, X., Cohen, K. J., Burger, P., Biegel, J. A., Kieran, M. W. (2009). Intensive multimodality treatment for children with newly diagnosed CNS atypical teratoid rhabdoid tumor. *J Clin Oncol*, 27(3), 385-389. doi: 10.1200/jco.2008.18.7724
- Chi, T. H. (2002). Reciprocal regulation of CD4/CD8 expression by SWI/SNF-like BAF complexes. *Nature*, 418, 195-199.
- Chiou, S. H., Kao, C. L., Chen, Y. W., Chien, C. S., Hung, S. C., Lo, J. F., Wong, T. T. (2008). Identification of CD133-positive radioresistant cells in atypical teratoid/rhabdoid tumor. *PLoS One*, 3(5), e2090. doi: 10.1371/journal.pone.0002090
- Cho, Y. M., Choi, J., Lee, O. J., Lee, H. I., Han, D. J., & Ro, J. Y. (2006). SMARCB1/INI1 missense mutation in mucinous carcinoma with rhabdoid features. *Pathol Int*, 56(11), 702-706. doi: 10.1111/j.1440-1827.2006.02033.x
- Chu, Y., & Corey, D. R. (2012). RNA sequencing: platform selection, experimental design, and data interpretation. *Nucleic Acid Ther*, 22(4), 271-274. doi: 10.1089/nat.2012.0367
- Chuang, J. C., & Jones, P. A. (2007). Epigenetics and microRNAs. *Pediatr Res*, 61(5 Pt 2), 24r-29r. doi: 10.1203/pdr.0b013e3180457684
- Chuang, T. H., Xu, X., Kaartinen, V., Heisterkamp, N., Groffen, J., & Bokoch, G. M. (1995). Abr and Bcr are multifunctional regulators of the Rho GTP-binding protein family. *Proc Natl Acad Sci U S A*, 92(22), 10282-10286.
- Churko, J. M., Mantalas, G. L., Snyder, M. P., & Wu, J. C. (2013). Overview of high throughput sequencing technologies to elucidate molecular pathways in cardiovascular diseases. *Circ Res*, 112(12), 1613-1623. doi: 10.1161/circresaha.113.300939
- Cohen, A. L., Holmen, S. L., & Colman, H. (2013). IDH1 and IDH2 mutations in gliomas. *Curr Neurol Neurosci Rep*, 13(5), 345. doi: 10.1007/s11910-013-0345-4

- Coisy-Quivy, M., Disson, O., Roure, V., Muchardt, C., Blanchard, J. M., & Dantonel, J. C. (2006). Role for Brm in cell growth control. *Cancer Res*, 66(10), 5069-5076. doi: 10.1158/0008-5472.can-05-0596
- Corona, D. F. V., & Tamkun, J. W. (2004). Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochimica et Biophysica Acta - Gene Structure and Expression*, 1677(1-3), 113-119.
- Cortez, C. C., & Jones, P. A. (2008). Chromatin, cancer and drug therapies. *Mutat Res*, 647(1-2), 44-51. doi: 10.1016/j.mrfmmm.2008.07.006
- Cottrell, C. E., Al-Kateb, H., Bredemeyer, A. J., Duncavage, E. J., Spencer, D. H., Abel, H. J., Pfeifer, J. D. (2014). Validation of a next-generation sequencing assay for clinical molecular oncology. *J Mol Diagn*, 16(1), 89-105. doi: 10.1016/j.jmoldx.2013.10.002
- Craig, E., Zhang, Z. K., Davies, K. P., & Kalpana, G. V. (2002). A masked NES in INI1/hSNF5 mediates hCRM1-dependent nuclear export: implications for tumorigenesis. *Embo j*, 21(1-2), 31-42. doi: 10.1093/emboj/21.1.31
- Dadmehr, M., Hosseini, M., Hosseinkhani, S., Ganjali, M. R., Khoobi, M., Behzadi, H., Sheikhnajad, R. (2014). DNA methylation detection by a novel fluorimetric nanobiosensor for early cancer diagnosis. *Biosensors and Bioelectronics*, 60, 35-44.
- Dammann, R., Schagdarsurengin, U., Strunnikova, M., Rastetter, M., Seidel, C., Liu, L., Pfeifer, G. P. (2003). Epigenetic inactivation of the Ras-association domain family 1 (RASSF1A) gene and its function in human carcinogenesis. *Histol Histopathol*, 18(2), 665-677.
- Dang, T., Vassilyadi, M., Michaud, J., Jimenez, C., & Ventureyra, E. C. (2003). Atypical teratoid/rhabdoid tumors. *Childs Nerv Syst*, 19(4), 244-248. doi: 10.1007/s00381-003-0731-3



- Darr, J., Klochendler, A., Isaac, S., & Eden, A. (2014). Loss of IGFBP7 expression and persistent AKT activation contribute to SMARCB1/Snf5-mediated tumorigenesis. *Oncogene*, 33(23), 3024-3032. doi: 10.1038/onc.2013.261
- Das, S., Banerjee, B., Hossain, M., Thangamuniyandi, M., Dasgupta, S., Chongdar, N., Basu, G. (2013). Characterization of DNA binding property of the HIV-1 host factor and tumor suppressor protein Integrase Interactor 1 (INI1/hSNF5). *PLoS One*, 8(7), e66581. doi: 10.1371/journal.pone.0066581
- David, R. B. (2005). *Child and adolescent neurology*. Wiley-Blackwell, 105.
- de Araujo, E. S., Marchi, F. A., Rodrigues, T. C., Vieira, H. C., Kuasne, H., Achatz, M. I., Krepischi, A. C. (2014). Genome-wide DNA methylation profile of leukocytes from melanoma patients with and without CDKN2A mutations. *Exp Mol Pathol*, 97(3), 425-432. doi: 10.1016/j.yexmp.2014.09.009
- De Carvalho, D. D., You, J. S., & Jones, P. A. (2010). DNA methylation and cellular reprogramming. *Trends Cell Biol*, 20(10), 609-617. doi: 10.1016/j.tcb.2010.08.003
- De Padua, M., Reddy, V., & Reddy, M. (2009). Cerebral atypical teratoid rhabdoid tumour arising in a child treated for acute lymphoblastic leukaemia. *BMJ Case Rep*, 2009. doi: 10.1136/bcr.08.2008.0601
- Decristofaro, M. F., Betz, B. L., Rorie, C. J., Reisman, D. N., Wang, W., & Weissman, B. E. (2001). Characterization of SWI/SNF protein expression in human breast cancer cell lines and other malignancies. *J Cell Physiol*, 186(1), 136-145. doi: 10.1002/1097-4652(200101)186:1<136::aid-jcp1010>3.0.co;2-4
- DeCristofaro, M. F., Betz, B. L., Wang, W., & Weissman, B. E. (1999). Alteration of hSNF5/INI1/BAF47 detected in rhabdoid cell lines and primary rhabdomyosarcomas but not Wilms' tumors. *Oncogene*, 18(52), 7559-7565. doi: 10.1038/sj.onc.1203168

- Dedeurwaerder, S., Defrance, M., Calonne, E., Denis, H., Sotiriou, C., & Fuks, F. (2011). Evaluation of the Infinium Methylation 450K technology. *Epigenomics*, 3(6), 771-784. doi: 10.2217/epi.11.105
- Deisch, J., Raisanen, J., & Rakheja, D. (2011). Immunohistochemical expression of embryonic stem cell markers in malignant rhabdoid tumors. *Pediatr Dev Pathol*, 14(5), 353-359. doi: 10.2350/10-09-0902-0a.1
- DelBove, J., Kuwahara, Y., Mora-Blanco, E. L., Godfrey, V., Funkhouser, W. K., Fletcher, C. D., Weissman, B. E. (2009). Inactivation of SNF5 cooperates with p53 loss to accelerate tumor formation in *Snf5*(+/-);*p53*(+/-) mice. *Mol Carcinog*, 48(12), 1139-1148. doi: 10.1002/mc.20568
- Dennis, G., Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*, 4(5), P3.
- Dennis, K., Fan, T., Geiman, T., Yan, Q., & Muegge, K. (2001). Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev*, 15(22), 2940-2944. doi: 10.1101/gad.929101
- Dennler, S., Andre, J., Alexaki, I., Li, A., Magnaldo, T., ten Dijke, P., Mauviel, A. (2007). Induction of sonic hedgehog mediators by transforming growth factor-beta: Smad3-dependent activation of Gli2 and Gli1 expression in vitro and in vivo. *Cancer Res*, 67(14), 6981-6986. doi: 10.1158/0008-5472.can-07-0491
- Dettling, M., & Buhlmann, P. (2002). Supervised clustering of genes. *Genome Biol*, 3(12), RESEARCH0069.
- Devarajan, K. (2008). Nonnegative matrix factorization: an analytical and interpretive tool in computational biology. *PLoS Comput Biol*, 4(7), e1000029. doi: 10.1371/journal.pcbi.1000029

Di Costanzo, A., Del Gaudio, N., Migliaccio, A., & Altucci, L. (2014). Epigenetic drugs against cancer: An evolving landscape. *Archives of Toxicology*, 88(9), 1651-1668.

Di Croce, L., & Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. *Nat Struct Mol Biol*, 20(10), 1147-1155. doi: 10.1038/nsmb.2669

Diaz-Rodriguez, E., Sotillo, R., Schwartzman, J. M., & Benezra, R. (2008). Hec1 overexpression hyperactivates the mitotic checkpoint and induces tumor formation in vivo. *Proc Natl Acad Sci U S A*, 105(43), 16719-16724. doi: 10.1073/pnas.0803504105

Doan, D. N., Veal, T. M., Yan, Z., Wang, W., Jones, S. N., & Imbalzano, A. N. (2004). Loss of the INI1 tumor suppressor does not impair the expression of multiple BRG1-dependent genes or the assembly of SWI/SNF enzymes. *Oncogene*, 23(19), 3462-3473. doi: 10.1038/sj.onc.1207472

Doan, D. N., Veal, T. M., Yan, Z., Wang, W., Jones, S. N., & Imbalzano, A. N. (2004). Loss of the INI1 tumor suppressor does not impair the expression of multiple BRG1-dependent genes or the assembly of SWI/SNF enzymes. *Oncogene*, 23(19), 3462-3473. doi: 10.1038/sj.onc.1207472

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15-21. doi: 10.1093/bioinformatics/bts635

Dong, W., Tu, S., Xie, J., Sun, P., Wu, Y., & Wang, L. (2009). Frequent promoter hypermethylation and transcriptional downregulation of BTG4 gene in gastric cancer. *Biochem Biophys Res Commun*, 387(1), 132-138. doi: <http://dx.doi.org/10.1016/j.bbrc.2009.06.140>

Drakaki, A., & Iliopoulos, D. (2013). MicroRNA-gene signaling pathways in pancreatic cancer. *Biomedical Journal*, 36(5), 200-208.

- Driemel, C., Kremling, H., Schumacher, S., Will, D., Wolters, J., Lindenlauf, N., Gires, O. (2014). Context-dependent adaption of EpCAM expression in early systemic esophageal cancer. *Oncogene*, 33(41), 4904-4915. doi: 10.1038/onc.2013.441
- Du, P., Kibbe, W. A., & Lin, S. M. (2008). lumi: a pipeline for processing Illumina microarray. *Bioinformatics*, 24(13), 1547-1548. doi: 10.1093/bioinformatics/btn224
- Du, P., Zhang, X., Huang, C. C., Jafari, N., Kibbe, W. A., Hou, L., & Lin, S. M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, 11, 587. doi: 10.1186/1471-2105-11-587
- Dufour, C., Beaugrand, A., Le Deley, M. C., Bourdeaut, F., Andre, N., Leblond, P., Varlet, P. (2012). Clinicopathologic prognostic factors in childhood atypical teratoid and rhabdoid tumor of the central nervous system: a multicenter study. *Cancer*, 118(15), 3812-3821. doi: 10.1002/cncr.26684
- Duhoux, F. P., Ameye, G., Montano-Almendras, C. P., Bahloula, K., Mozziconacci, M. J., Laibe, S., Poirel, H. A. (2012). PRDM16 (1p36) translocations define a distinct entity of myeloid malignancies with poor prognosis but may also occur in lymphoid malignancies. *Br J Haematol*, 156(1), 76-88. doi: 10.1111/j.1365-2141.2011.08918.x
- Dunaief, J. L., Strober, B. E., Guha, S., Khavari, P. A., Alin, K., Luban, J., Goff, S. P. (1994). The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. *Cell*, 79(1), 119-130.
- Dunkel, I. J., Boyett, J. M., Yates, A., Rosenblum, M., Garvin, J. H., Jr., Bostrom, B. C., Finlay, J. L. (1998). High-dose carboplatin, thiotepa, and etoposide with autologous stem-cell rescue for patients with recurrent medulloblastoma. Children's Cancer Group. *J Clin Oncol*, 16(1), 222-228.
- Dunning, M. J., Smith, M. L., Ritchie, M. E., & Tavaré, S. (2007). beadarray: R classes and methods for Illumina bead-based data. *Bioinformatics*, 23(16), 2183-2184. doi: 10.1093/bioinformatics/btm311

- Duxin, J. P., Dao, B., Martinsson, P., Rajala, N., Guittat, L., Campbell, J. L., Stewart, S. A. (2009). Human Dna2 is a nuclear and mitochondrial DNA maintenance protein. *Mol Cell Biol*, 29(15), 4274-4282. doi: 10.1128/mcb.01834-08
- Eaton, K. W., Tooke, L. S., Wainwright, L. M., Judkins, A. R., & Biegel, J. A. (2011). Spectrum of SMARCB1/INI1 mutations in familial and sporadic rhabdoid tumors. *Pediatr Blood Cancer*, 56(1), 7-15. doi: 10.1002/pbc.22831
- Eberhart, C. G. (2011). Molecular diagnostics in embryonal brain tumors. *Brain Pathol*, 21(1), 96-104. doi: 10.1111/j.1750-3639.2010.00455.x
- Eden, A., Gaudet, F., Waghmare, A., & Jaenisch, R. (2003). Chromosomal Instability and Tumors Promoted by DNA Hypomethylation. *Science*, 300(5618), 455. doi: 10.1126/science.1083557
- Edgren, H., Murumagi, A., Kangaspeska, S., Nicorici, D., Hongisto, V., Kleivi, K., Kallioniemi, O. (2011). Identification of fusion genes in breast cancer by paired-end RNA-sequencing. *Genome Biol*, 12(1), R6. doi: 10.1186/gb-2011-12-1-r6
- Ehrlich, M. (2009). DNA hypomethylation in cancer cells. *Epigenomics*, 1(2), 239-259. doi: 10.2217/epi.09.33
- Ehrlich, M., Woods, C. B., Yu, M. C., Dubeau, L., Yang, F., Campan, M., Laird, P. W. (2006). Quantitative analysis of associations between DNA hypermethylation, hypomethylation, and DNMT RNA levels in ovarian tumors. *Oncogene*, 25(18), 2636-2645. doi: 10.1038/sj.onc.1209145
- Elkouris, M., Balaskas, N., Poulou, M., Politis, P. K., Panayiotou, E., Malas, S., Remboutsika, E. (2011). Sox1 maintains the undifferentiated state of cortical neural progenitor cells via the suppression of Prox1-mediated cell cycle exit and neurogenesis. *Stem Cells*, 29(1), 89-98. doi: 10.1002/stem.554

- Elkouris, M., Balaskas, N., Poulou, M., Politis, P. K., Panayiotou, E., Malas, S., Remboutsika, E. (2011). Sox1 maintains the undifferentiated state of cortical neural progenitor cells via the suppression of Prox1-mediated cell cycle exit and neurogenesis. *Stem Cells*, 29(1), 89-98. doi: 10.1002/stem.554
- Elmore, S. (2007). Apoptosis: a review of programmed cell death. *Toxicol Pathol*, 35(4), 495-516. doi: 10.1080/01926230701320337
- Engel, M., Seifert, M., Theisinger, B., Seyfert, U., & Welter, C. (1998). Glyceraldehyde-3-phosphate dehydrogenase and Nm23-H1/nucleoside diphosphate kinase A. Two old enzymes combine for the novel Nm23 protein phosphotransferase function. *J Biol Chem*, 273(32), 20058-20065.
- Esteller, M. (2002). CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene*, 21(35), 5427-5440. doi: 10.1038/sj.onc.1205600
- Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet*, 8(4), 286-298.
- Esteller, M. (2008). Epigenetics in Cancer. *New England Journal of Medicine*, 358(11), 1148-1159. doi: doi:10.1056/NEJMra072067
- Euskirchen, G., Auerbach, R. K., & Snyder, M. (2012). SWI/SNF chromatin-remodeling factors: multiscale analyses and diverse functions. *J Biol Chem*, 287(37), 30897-30905. doi: 10.1074/jbc.R111.309302
- Euskirchen, G. M., Auerbach, R. K., Davidov, E., Gianoulis, T. A., Zhong, G., Rozowsky, J., Snyder, M. (2011). Diverse roles and interactions of the SWI/SNF chromatin remodeling complex revealed using global approaches. *PLoS Genet*, 7(3), e1002008. doi: 10.1371/journal.pgen.1002008

- Eveno, C., Contreres, J. O., Hainaud, P., Nemeth, J., Dupuy, E., & Pocard, M. (2013). Netrin-4 overexpression suppresses primary and metastatic colorectal tumor progression. *Oncol Rep*, 29(1), 73-78. doi: 10.3892/or.2012.2104
- Fanelli, M., Amatori, S., Barozzi, I., Soncini, M., Dal Zuffo, R., Bucci, G., Minucci, S. (2010). Pathology tissue-chromatin immunoprecipitation, coupled with high-throughput sequencing, allows the epigenetic profiling of patient samples. *Proc Natl Acad Sci U S A*, 107(50), 21535-21540. doi: 10.1073/pnas.1007647107
- Fedele, M., Visone, R., De Martino, I., Troncone, G., Palmieri, D., Battista, S., Fusco, A. (2006). HMGA2 induces pituitary tumorigenesis by enhancing E2F1 activity. *Cancer Cell*, 9(6), 459-471. doi: 10.1016/j.ccr.2006.04.024
- Ferrara, G., Pancione, M., Votino, C., Quaglino, P., Tomasini, C., Santucci, M., Colantuoni, V. (2014). A specific DNA methylation profile correlates with a high risk of disease progression in stage i classical (Alibert-Bazin type) mycosis fungoides. *British Journal of Dermatology*, 170(6), 1266-1275.
- Fidani, P., De Ioris, M. A., Serra, A., De Sio, L., Ilari, I., Cozza, R., Donfrancesco, A. (2009). A multimodal strategy based on surgery, radiotherapy, ICE regimen and high dose chemotherapy in atypical teratoid/rhabdoid tumours: a single institution experience. *J Neurooncol*, 92(2), 177-183. doi: 10.1007/s11060-008-9750-y
- Finetti M. A. , D. C. P. A., Selby M., Smith A., Crosier S., Wood J., Skalkoyannis B., Bailey S., Clifford S., Williamson D. (2014). Characterising critical pathways in ATRT tumourigenesis: a genome-wide analysis of primary tumours and functional genomic models. *Neuro Oncol*(16th International Symposium on Pediatric Neuro-Oncology (ISPNO) in conjunction with the 8th St. Jude-VIVA Forum).
- Finetti M. A. , D. C. P. A., Wood J., Skalkoyannis B., Selby M., Smith A., Crosier S., Bailey S., Clifford S., Williamson D. (2014). Mechanism of tumourigenesis caused by loss of SMARCB1 in malignant rhabdoid tumors. *European Journal of Cancer*, 50, Supplement 5(EACR-23: From Basic Research to Personalised Cancer Treatment), 470.

Finetti M. A. , D. C. P. A., Skalkoyannis B., Selby M., Smith A., Crosier S., Bailey S., Clifford S., Williamson D. (2014). A Systems Biology Approach to Identify Mechanism of Tumourigenesis Caused by Loss of SMARCB1 in Malignant Rhabdoid Tumors. (2014 NCRI cancer conference).

Finetti M. A. , S. M., Del Carpio Pons A, Wood J., Skalkoyannis B., , Smith A., Crosier S., Bailey S., Clifford S., Williamson D. (2014). Next-generation sequencing identifies the mechanism of tumourigenesis caused by loss of SMARCB1 in malignant rhabdoid tumours. *European Journal of Cancer*, 50, supplement 6(Conference: 26 EORTC – NCI – AACR Symposium on th Molecular Targets and Cancer Therapeutics), 413. doi: 10.1016/S0959-8049(14)70539-8

Flaus, A., & Owen-Hughes, T. (2004). Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer? *Curr Opin Genet Dev*, 14(2), 165-173. doi: 10.1016/j.gde.2004.01.007

Fog, C. K., Jensen, K. T., & Lund, A. H. (2007). Chromatin-modifying proteins in cancer. *Apmis*, 115(10), 1060-1089. doi: 10.1111/j.1600-0463.2007.apm\_776.xml.x

Forbes, N. E., Abdelbary, H., Lupien, M., Bell, J. C., & Diallo, J. S. (2013). Exploiting tumor epigenetics to improve oncolytic virotherapy. *Frontiers in Genetics*, 4(SEP).

Forest, F., David, A., Arrufat, S., Pierron, G., Ranchere-Vince, D., Stephan, J. L., Bourdeaut, F. (2012). Conventional chondrosarcoma in a survivor of rhabdoid tumor: enlarging the spectrum of tumors associated with SMARCB1 germline mutations. *Am J Surg Pathol*, 36(12), 1892-1896. doi: 10.1097/PAS.0b013e31826cbe7a

Foster, K. S., McCrary, W. J., Ross, J. S., & Wright, C. F. (2006). Members of the hSWI/SNF chromatin remodeling complex associate with and are phosphorylated by protein kinase B/Akt. *Oncogene*, 25(33), 4605-4612. doi: 10.1038/sj.onc.1209496

Franckevica, I., Kleina, R., & Voika, O. (2011). Originally misdiagnosed rhabdoid tumour of the kidney. A case report and differential diagnosis. *Pol J Pathol*, 62(3), 163-167.



Fruhwald, M. C., Hasselblatt, M., Wirth, S., Kohler, G., Schneppenheim, R., Subero, J. I., Vormoor, J. (2006). Non-linkage of familial rhabdoid tumors to SMARCB1 implies a second locus for the rhabdoid tumor predisposition syndrome. *Pediatr Blood Cancer*, 47(3), 273-278. doi: 10.1002/pbc.20526

Fujisawa, H., Misaki, K., Takabatake, Y., Hasegawa, M., & Yamashita, J. (2005). Cyclin D1 is overexpressed in atypical teratoid/rhabdoid tumor with hSNF5/INI1 gene inactivation. *J Neurooncol*, 73(2), 117-124. doi: 10.1007/s11060-004-4276-4

Fujita, M., Sato, M., Nakamura, M., Kudo, K., Nagasaka, T., Mizuno, M., Yoshida, J. (2005). Multicentric atypical teratoid/rhabdoid tumors occurring in the eye and fourth ventricle of an infant: case report. *J Neurosurg*, 102(3 Suppl), 299-302. doi: 10.3171/ped.2005.102.3.0299

Gadd, S., Sredni, S. T., Huang, C.-C., & Perlman, E. J. (2010). Rhabdoid tumor: gene expression clues to pathogenesis and potential therapeutic targets. *Lab Invest*, 90(5), 724-738. doi: <http://www.nature.com/labinvest/journal/v90/n5/supinfo/labinvest201066s1.html>

Gal-Yam, E. N., Saito, Y., Egger, G., & Jones, P. A. (2008). Cancer epigenetics: modifications, screening, and therapy. *Annu Rev Med*, 59, 267-280. doi: 10.1146/annurev.med.59.061606.095816

Ganger, D. R., Hamilton, P. D., Fletcher, J. W., & Fernandez-Pol, J. A. (1997). Metallopanstimulin is overexpressed in a patient with colonic carcinoma. *Anticancer Res*, 17(3C), 1993-1999.

Ganger, D. R., Hamilton, P. D., Klos, D. J., Jakate, S., McChesney, L., & Fernandez-Pol, J. A. (2001). Differential expression of metallopanstimulin/S27 ribosomal protein in hepatic regeneration and neoplasia. *Cancer Detect Prev*, 25(3), 231-236.

- Garber, M., Grabherr, M. G., Guttman, M., & Trapnell, C. (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Meth*, 8(6), 469-477. doi: <http://www.nature.com/nmeth/journal/v8/n6/abs/nmeth.1613.html#supplementary-information>
- Garre, M. L., & Tekautz, T. (2010). Role of high-dose chemotherapy (HDCT) in treatment of atypical teratoid/rhabdoid tumors (AT/RTs). *Pediatr Blood Cancer*, 54(4), 647-648. doi: 10.1002/pbc.22377
- Gaziel-Sovran, A., Segura, M. F., Di Micco, R., Collins, M. K., Hanniford, D., Vega-Saenz de Miera, E., Hernando, E. (2011). miR-30b/30d regulation of GalNAc transferases enhances invasion and immunosuppression during metastasis. *Cancer Cell*, 20(1), 104-118. doi: 10.1016/j.ccr.2011.05.027
- Gérard, C., Gonze, D., Lemaigre, F., & Novák, B. (2014). A Model for the Epigenetic Switch Linking Inflammation to Cell Transformation: Deterministic and Stochastic Approaches. *PLoS Computational Biology*, 10(1).
- Gidwani, P., Levy, A., Goodrich, J., Weidenheim, K., & Kolb, E. A. (2008). Successful outcome with tandem myeloablative chemotherapy and autologous peripheral blood stem cell transplants in a patient with atypical teratoid/rhabdoid tumor of the central nervous system. *J Neurooncol*, 88(2), 211-215. doi: 10.1007/s11060-008-9553-1
- Gilchrist, D. A., Fargo, D. C., & Adelman, K. (2009). Using ChIP-chip and ChIP-seq to study the regulation of gene expression: genome-wide localization studies reveal widespread regulation of transcription elongation. *Methods*, 48(4), 398-408. doi: 10.1016/j.ymeth.2009.02.024
- Ginn, K. F., & Gajjar, A. (2012). Atypical teratoid rhabdoid tumor: current therapy and future directions. *Front Oncol*, 2, 114. doi: 10.3389/fonc.2012.00114

- Giraud, S., Hurlstone, A., Avril, S., & Coqueret, O. (2004). Implication of BRG1 and cdk9 in the STAT3-mediated activation of the p21waf1 gene. *Oncogene*, 23(44), 7391-7398. doi: 10.1038/sj.onc.1207972
- Glaros, S., Cirrincione, G. M., Muchardt, C., Kleer, C. G., Michael, C. W., & Reisman, D. (2007). The reversible epigenetic silencing of BRM: implications for clinical targeted therapy. *Oncogene*, 26(49), 7058-7066. doi: 10.1038/sj.onc.1210514
- Godbout, R., Bisgrove, D. A., Shkolny, D., & Day, R. S., 3rd. (1998). Correlation of B-FABP and GFAP expression in malignant glioma. *Oncogene*, 16(15), 1955-1962. doi: 10.1038/sj.onc.1201740
- Gong, F., Fahy, D., & Smerdon, M. J. (2006). Rad4-Rad23 interaction with SWI/SNF links ATP-dependent chromatin remodeling with nucleotide excision repair. *Nat Struct Mol Biol*, 13(10), 902-907. doi: 10.1038/nsmb1152
- Gorayski, P., Boros, S., Ong, B., Olson, S., & Foote, M. (2013). Radiation-induced primary cerebral atypical teratoid/rhabdoid tumour in an adult. *J Clin Neurosci*, 20(10), 1466-1468. doi: 10.1016/j.jocn.2013.03.041
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotech*, 29(7), 644-652. doi: <http://www.nature.com/nbt/journal/v29/n7/abs/nbt.1883.html#supplementary-information>
- Grada, A., & Weinbrecht, K. (2013). Next-Generation Sequencing: Methodology and Application. *J Invest Dermatol*, 133(8), e11. doi: 10.1038/jid.2013.248
- Gramling, S., & Reisman, D. (2011). Discovery of BRM Targeted Therapies: Novel Reactivation of an Anti-cancer Gene. *Lett Drug Des Discov*, 8(1), 93-99.

- Gramling, S., Rogers, C., Liu, G., & Reisman, D. (2011). Pharmacologic reversal of epigenetic silencing of the anticancer protein BRM: a novel targeted treatment strategy. *Oncogene*, 30(29), 3289-3294. doi: 10.1038/onc.2011.80
- Grand, F., Kulkarni, S., Chase, A., Goldman, J. M., Gordon, M., & Cross, N. C. (1999). Frequent deletion of hSNF5/INI1, a component of the SWI/SNF complex, in chronic myeloid leukemia. *Cancer Res*, 59(16), 3870-3874.
- Gresh, L., Bourachot, B., Reimann, A., Guigas, B., Fiette, L., Garbay, S., Klochendler-Yeivin, A. (2005). The SWI/SNF chromatin-remodeling complex subunit SNF5 is essential for hepatocyte differentiation. *Embo j*, 24(18), 3313-3324. doi: 10.1038/sj.emboj.7600802
- Gronbaek, K., Hother, C., & Jones, P. A. (2007). Epigenetic changes in cancer. *Apmis*, 115(10), 1039-1059. doi: 10.1111/j.1600-0463.2007.apm\_636.xml.x
- Guan, M., Pang, C.-P., Yam, H.-F., Cheung, K.-F., Liu, W.-W., & Lu, Y. (2004). Inhibition of glioma invasion by overexpression of pigment epithelium-derived factor. *Cancer Gene Ther*, 11(5), 325-332.
- Guidi, C. J., Mudhasani, R., Hoover, K., Koff, A., Leav, I., Imbalzano, A. N., & Jones, S. N. (2006). Functional interaction of the retinoblastoma and *Ini1/Snf5* tumor suppressors in cell growth and pituitary tumorigenesis. *Cancer Res*, 66(16), 8076-8082. doi: 10.1158/0008-5472.can-06-1451
- Guidi, C. J., Sands, A. T., Zambrowicz, B. P., Turner, T. K., Demers, D. A., Webster, W., Jones, S. N. (2001). Disruption of *Ini1* leads to peri-implantation lethality and tumorigenesis in mice. *Mol Cell Biol*, 21(10), 3598-3603. doi: 10.1128/mcb.21.10.3598-3603.2001
- Guidi, C. J., Veal, T. M., Jones, S. N., & Imbalzano, A. N. (2004). Transcriptional compensation for loss of an allele of the *Ini1* tumor suppressor. *J Biol Chem*, 279(6), 4180-4185. doi: 10.1074/jbc.M312043200

Gunawardena, R. W., Siddiqui, H., Solomon, D. A., Mayhew, C. N., Held, J., Angus, S. P., & Knudsen, E. S. (2004). Hierarchical requirement of SWI/SNF in retinoblastoma tumor suppressor-mediated repression of Plk1. *J Biol Chem*, 279(28), 29278-29285. doi: 10.1074/jbc.M400395200

Guo, X., Li, X., Wang, Y., Tian, Z., Duan, X., & Cai, Z. (2014). Nicotine induces alteration of H3K27 demethylase UTX in kidney cancer cell. *Human and Experimental Toxicology*, 33(3), 264-269.

Gupta, R., Wikramasinghe, P., Bhattacharyya, A., Perez, F. A., Pal, S., & Davuluri, R. V. (2010). Annotation of gene promoters by integrative data-mining of ChIP-seq Pol-II enrichment data. *BMC Bioinformatics*, 11 Suppl 1, S65. doi: 10.1186/1471-2105-11-s1-s65

Gutierrez, J., Paredes, R., Cruzat, F., Hill, D. A., van Wijnen, A. J., Lian, J. B., Montecino, M. (2007). Chromatin remodeling by SWI/SNF results in nucleosome mobilization to preferential positions in the rat osteocalcin gene promoter. *J Biol Chem*, 282(13), 9445-9457. doi: 10.1074/jbc.M609847200

Haberler, C. (2012). Intact INI1 gene region with paradoxical loss of protein expression in AT/RT: implications for a possible novel mechanism associated with absence of INI1 protein immunoreactivity. *Am J Surg Pathol*, 36(12), 1903; author reply 1903-1904. doi: 10.1097/PAS.0b013e31825d53d5

Haberler, C., Laggner, U., Slavec, I., Czech, T., Ambros, I. M., Ambros, P. F., Hainfellner, J. A. (2006). Immunohistochemical analysis of INI1 protein in malignant pediatric CNS tumors: Lack of INI1 in atypical teratoid/rhabdoid tumors and in a fraction of primitive neuroectodermal tumors without rhabdoid phenotype. *Am J Surg Pathol*, 30(11), 1462-1468. doi: 10.1097/01.pas.0000213329.71745.ef

- Haberler, C., Laggner, U., Slavc, I., Czech, T., Ambros, I. M., Ambros, P. F., Hainfellner, J. A. (2006). Immunohistochemical analysis of INI1 protein in malignant pediatric CNS tumors: Lack of INI1 in atypical teratoid/rhabdoid tumors and in a fraction of primitive neuroectodermal tumors without rhabdoid phenotype. *Am J Surg Pathol*, 30(11), 1462-1468. doi: 10.1097/01.pas.0000213329.71745.ef
- Hah, N., Kolkman, A., Ruhl, D. D., Pijnappel, W. W., Heck, A. J., Timmers, H. T., & Kraus, W. L. (2010). A role for BAF57 in cell cycle-dependent transcriptional regulation by the SWI/SNF chromatin remodeling complex. *Cancer Res*, 70(11), 4402-4411. doi: 10.1158/0008-5472.can-09-2767
- Hahn, M. A., Hahn, T., Lee, D. H., Esworthy, R. S., Kim, B. W., Riggs, A. D., Pfeifer, G. P. (2008). Methylation of polycomb target genes in intestinal cancer is mediated by inflammation. *Cancer Res*, 68(24), 10280-10289. doi: 10.1158/0008-5472.can-08-1957
- Halliday, G. M., Bock, V. L., Moloney, F. J., & Lyons, J. G. (2009). SWI/SNF: a chromatin-remodelling complex with a role in carcinogenesis. *Int J Biochem Cell Biol*, 41(4), 725-728. doi: 10.1016/j.biocel.2008.04.026
- Hama, R., Watanabe, Y., Shinada, K., Yamada, Y., Ogata, Y., Yoshida, Y., Itoh, F. (2012). Characterization of DNA hypermethylation in two cases of peritoneal mesothelioma. *Tumour Biol*, 33(6), 2031-2040. doi: 10.1007/s13277-012-0462-8
- Hansen, K. D., Brenner, S. E., & Dudoit, S. (2010). Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res*, 38(12), e131. doi: 10.1093/nar/gkq224
- Harada, K., Toyooka, S., Maitra, A., Maruyama, R., Toyooka, K. O., Timmons, C. F., Gazdar, A. F. (2002). Aberrant promoter methylation and silencing of the RASSF1A gene in pediatric tumors and cell lines. *Oncogene*, 21(27), 4345-4349. doi: 10.1038/sj.onc.1205446
- Hasan, A., Palumbo, M., Atkinson, J., Carret, A. S., Farmer, J. P., Montes, J., Freeman, C. R. (2011). Treatment-related morbidity in atypical teratoid/rhabdoid tumor:

multifocal necrotizing leukoencephalopathy. *Pediatr Neurosurg*, 47(1), 7-14. doi: 10.1159/000323412

Hashimoto, K., Kokubun, S., Itoi, E., & Roach, H. I. (2007). Improved quantification of DNA methylation using methylation-sensitive restriction enzymes and real-time PCR. *Epigenetics*, 2(2), 86-91.

Hasselblatt, M., Gesk, S., Oyen, F., Rossi, S., Viscardi, E., Giangaspero, F., Paulus, W. (2011). Nonsense mutation and inactivation of SMARCA4 (BRG1) in an atypical teratoid/rhabdoid tumor showing retained SMARCB1 (INI1) expression. *Am J Surg Pathol*, 35(6), 933-935. doi: 10.1097/PAS.0b013e3182196a39

Hasselblatt, M., Isken, S., Linge, A., Eikmeier, K., Jeibmann, A., Oyen, F., Paulus, W. (2013). High-resolution genomic analysis suggests the absence of recurrent genomic alterations other than SMARCB1 aberrations in atypical teratoid/rhabdoid tumors. *Genes Chromosomes Cancer*, 52(2), 185-190. doi: 10.1002/gcc.22018

Hatzia Apostolou, M., & Iliopoulos, D. (2011). Epigenetic aberrations during oncogenesis. *Cellular and Molecular Life Sciences*, 68(10), 1681-1702. doi: 10.1007/s00018-010-0624-z

He, H., & Sun, Y. (2006). Ribosomal protein S27L is a direct p53 target that regulates apoptosis. *Oncogene*, 26(19), 2707-2716. doi: <http://www.nature.com/onc/journal/v26/n19/supinfo/1210073s1.html>

He, S., Pirity, M. K., Wang, W. L., Wolf, L., Chauhan, B. K., Cveklova, K., Cvekl, A. (2010). Chromatin remodeling enzyme Brg1 is required for mouse lens fiber cell terminal differentiation and its denucleation. *Epigenetics Chromatin*, 3(1), 21. doi: 10.1186/1756-8935-3-21

He, Y. J., Zhang, Z. D., Yin, M. Z., & Wu, X. R. (2012). [Atypical teratoid/rhabdoid tumors of central nervous system in childhood: a clinical and histopathologic study of 6 cases]. *Zhonghua Bing Li Xue Za Zhi*, 41(4), 220-223. doi: 10.3760/cma.j.issn.0529-5807.2012.04.002

- Heck, J. E., Lombardi, C. A., Cockburn, M., Meyers, T. J., Wilhelm, M., & Ritz, B. (2013). Epidemiology of rhabdoid tumors of early childhood. *Pediatr Blood Cancer*, 60(1), 77-81. doi: 10.1002/pbc.24141
- Helming, K. C., Wang, X., & Roberts, C. W. (2014). Vulnerabilities of Mutant SWI/SNF Complexes in Cancer. *Cancer Cell*, 26(3), 309-317. doi: 10.1016/j.ccr.2014.07.018
- Helmke, L., Engler, S., Mattke, A., & Henne-Bruns, D. (2001). Extrarenal malignant rhabdoid tumors in childhood. *Med Pediatr Oncol*, 36(2), 317-319. doi: 10.1002/1096-911x(20010201)36:2<317::aid-mpo1073>3.0.co;2-w
- Hepp, M. I., Alarcon, V., Dutta, A., Workman, J. L., & Gutierrez, J. L. (2014). Nucleosome remodeling by the SWI/SNF complex is enhanced by yeast High Mobility Group Box (HMGB) proteins. *Biochim Biophys Acta*, 1839(9), 764-772. doi: 10.1016/j.bbagr.2014.06.014
- Herman, J. G., & Baylin, S. B. (2003). Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med*, 349(21), 2042-2054. doi: 10.1056/NEJMr023075
- Hertwig, F., Meyer, K., Braun, S., Ek, S., Spang, R., Pfenninger, C. V., Nuber, U. A. (2012). Definition of genetic events directing the development of distinct types of brain tumors from postnatal neural stem/progenitor cells. *Cancer Res*, 72(13), 3381-3392. doi: 10.1158/0008-5472.can-11-3525
- Heuer, G. G., Kiefer, H., Judkins, A. R., Belasco, J., Biegel, J. A., Jackson, E. M., Storm, P. B. (2010). Surgical treatment of a clival-C2 atypical teratoid/rhabdoid tumor. *J Neurosurg Pediatr*, 5(1), 75-79. doi: 10.3171/2009.8.peds08421
- Higashino, K., Narita, T., Taga, T., Ohta, S., & Takeuchi, Y. (2003). Malignant rhabdoid tumor shows a unique neural differentiation as distinct from neuroblastoma. *Cancer Sci*, 94(1), 37-42.



Hilden, J. M., Meerbaum, S., Burger, P., Finlay, J., Janss, A., Scheithauer, B. W., Biegel, J. A. (2004). Central nervous system atypical teratoid/rhabdoid tumor: results of therapy in children enrolled in a registry. *J Clin Oncol*, 22(14), 2877-2884. doi: 10.1200/jco.2004.07.073

Ho, D. M., Hsu, C. Y., Wong, T. T., Ting, L. T., & Chiang, H. (2000). Atypical teratoid/rhabdoid tumor of the central nervous system: a comparative study with primitive neuroectodermal tumor/medulloblastoma. *Acta Neuropathol*, 99(5), 482-488.

Ho, L. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. *Proc. Natl Acad. Sci. USA*, 106, 5187-5191.

Ho, L. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc. Natl Acad. Sci. USA*, 106, 5181-5186.

Ho, L., & Crabtree, G. R. (2010). Chromatin remodelling during development. *Nature*, 463(7280), 474-484.

Hoell, J. I., Gombert, M., Bartenhagen, C., Ginzel, S., Husemann, P., Felsberg, J., Fischer, U. (2013). Whole-genome paired-end analysis confirms remarkable genomic stability of atypical teratoid/rhabdoid tumors. *Genes Chromosomes Cancer*, 52(10), 983-985. doi: 10.1002/gcc.22092

Hoffman, G. R., Rahal, R., Buxton, F., Xiang, K., McAllister, G., Frias, E., Jagani, Z. (2014). Functional epigenetics approach identifies BRM/SMARCA2 as a critical synthetic lethal target in BRG1-deficient cancers. *Proc Natl Acad Sci U S A*, 111(8), 3128-3133. doi: 10.1073/pnas.1316793111

Hohmann, A. F., & Vakoc, C. R. (2014). A rationale to target the SWI/SNF complex for cancer therapy. *Trends Genet*, 30(8), 356-363. doi: 10.1016/j.tig.2014.05.001

- Hollmann, T. J., & Hornick, J. L. (2011). INI1-deficient tumors: diagnostic features and molecular genetics. *Am J Surg Pathol*, 35(10), e47-63. doi: 10.1097/PAS.0b013e31822b325b
- Hovestadt, V., Jones, D. T., Picelli, S., Wang, W., Kool, M., Northcott, P. A., Lichter, P. (2014). Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature*, 510(7506), 537-541. doi: 10.1038/nature13268
- Howard, G., Eiges, R., Gaudet, F., Jaenisch, R., & Eden, A. (2008). Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene*, 27(3), 404-408. doi: 10.1038/sj.onc.1210631
- Hsu, C. C., Lee, Y. C., Yeh, S. H., Chen, C. H., Wu, C. C., Wang, T. Y., Lin, D. Y. (2012). 58-kDa microspherule protein (MSP58) is novel brahma-related gene 1 (BRG1)-associated protein that modulates p53/p21 senescence pathway. *Journal of Biological Chemistry*, 287(27), 22533-22548.
- Hu, F., Gartenhaus, R. B., Eichberg, D., Liu, Z., Fang, H. B., & Rapoport, A. P. (2010). PBK/TOPK interacts with the DBD domain of tumor suppressor p53 and modulates expression of transcriptional targets including p21. *Oncogene*, 29(40), 5464-5474. doi: <http://www.nature.com/onc/journal/v29/n40/supinfo/onc2010275s1.html>
- Hu, Y., Ylivinkka, I., Chen, P., Li, L., Hautaniemi, S., Nyman, T. A., Hyytiainen, M. (2012). Netrin-4 promotes glioblastoma cell proliferation through integrin beta4 signaling. *Neoplasia*, 14(3), 219-227.
- Hulsebos, T. J., Kenter, S., Verhagen, W. I., Baas, F., Flucke, U., & Wesseling, P. (2014). Premature termination of SMARCB1 translation may be followed by reinitiation in schwannomatosis-associated schwannomas, but results in absence of SMARCB1 expression in rhabdoid tumors. *Acta Neuropathol*, 128(3), 439-448. doi: 10.1007/s00401-014-1281-3

Hulsebos, T. J., Kenter, S. B., Jakobs, M. E., Baas, F., Chong, B., & Delatycki, M. B. (2010). SMARCB1/INI1 maternal germ line mosaicism in schwannomatosis. *Clin Genet*, 77(1), 86-91. doi: 10.1111/j.1399-0004.2009.01249.x

Hulsebos, T. J., Plomp, A. S., Wolterman, R. A., Robanus-Maandag, E. C., Baas, F., & Wesseling, P. (2007). Germline mutation of INI1/SMARCB1 in familial schwannomatosis. *Am J Hum Genet*, 80(4), 805-810. doi: 10.1086/513207

Ilyas, A., Ahmad, S., Faheem, M., Naseer, M., Kumosani, T., Al-Qahtani, M., Ahmed, F. (2015). Next Generation Sequencing of Acute Myeloid Leukemia: Influencing Prognosis. *BMC Genomics*, 16(Suppl 1), S5.

Inoue, N., Watanabe, H., Okamura, K., Sakaki, M., Kageji, T., Nagahiro, S., & Kagami, S. (2014). Atypical teratoid rhabdoid tumor in the cavernous sinus of a toddler presenting with oculomotor nerve palsy. *Childs Nerv Syst*, 30(8), 1463-1466. doi: 10.1007/s00381-014-2407-6

Institute Curie , P. (unpublished).

Isaacs, H., Jr. (2008). Fetal and neonatal renal tumors. *J Pediatr Surg*, 43(9), 1587-1595. doi: 10.1016/j.jpedsurg.2008.03.052

Isaacs, H., Jr. (2010). Fetal and neonatal rhabdoid tumor. *J Pediatr Surg*, 45(3), 619-626. doi: 10.1016/j.jpedsurg.2009.12.011

Isakoff, M. S., Sansam, C. G., Tamayo, P., Subramanian, A., Evans, J. A., Fillmore, C. M., Roberts, C. W. (2005). Inactivation of the Snf5 tumor suppressor stimulates cell cycle progression and cooperates with p53 loss in oncogenic transformation. *Proc Natl Acad Sci U S A*, 102(49), 17745-17750. doi: 10.1073/pnas.0509014102

Ito, T., Yamauchi, M., Nishina, M., Yamamichi, N., Mizutani, T., Ui, M., Iba, H. (2001). Identification of SWI.SNF complex subunit BAF60a as a determinant of the transactivation potential of Fos/Jun dimers. *J Biol Chem*, 276(4), 2852-2857. doi: 10.1074/jbc.M009633200

Iwasaki, H., Nakano, K., Shinkai, K., Kunisawa, Y., Hirahashi, M., Oda, Y., Katano, M. (2013). Hedgehog Gli3 activator signal augments tumorigenicity of colorectal cancer via upregulation of adherence-related genes. *Cancer Sci*, 104(3), 328-336. doi: 10.1111/cas.12073

Jackson, E. M., Shaikh, T. H., Gururangan, S., Jones, M. C., Malkin, D., Nikkel, S. M., Biegel, J. A. (2007). High-density single nucleotide polymorphism array analysis in patients with germline deletions of 22q11.2 and malignant rhabdoid tumor. *Hum Genet*, 122(2), 117-127. doi: 10.1007/s00439-007-0386-3

Jackson, E. M., Sievert, A. J., Gai, X., Hakonarson, H., Judkins, A. R., Tooke, L., Biegel, J. A. (2009). Genomic analysis using high-density single nucleotide polymorphism-based oligonucleotide arrays and multiplex ligation-dependent probe amplification provides a comprehensive analysis of INI1/SMARCB1 in malignant rhabdoid tumors. *Clin Cancer Res*, 15(6), 1923-1930. doi: 10.1158/1078-0432.ccr-08-2091

Jackson, E. M., Sievert, A. J., Gai, X., Hakonarson, H., Judkins, A. R., Tooke, L., Biegel, J. A. (2009). Genomic analysis using high-density single nucleotide polymorphism-based oligonucleotide arrays and multiplex ligation-dependent probe amplification provides a comprehensive analysis of INI1/SMARCB1 in malignant rhabdoid tumors. *Clin Cancer Res*, 15(6), 1923-1930. doi: 10.1158/1078-0432.ccr-08-2091

Jagani, Z., Mora-Blanco, E. L., Sansam, C. G., McKenna, E. S., Wilson, B., Chen, D., Dorsch, M. (2010). Loss of the tumor suppressor Snf5 leads to aberrant activation of the Hedgehog-Gli pathway. *Nat Med*, 16(12), 1429-1433. doi: 10.1038/nm.2251

Jagani, Z., Mora-Blanco, E. L., Sansam, C. G., McKenna, E. S., Wilson, B., Chen, D., Dorsch, M. (2010). Loss of the tumor suppressor Snf5 leads to aberrant activation of the Hedgehog-Gli pathway. *Nat Med*, 16(12), 1429-1433. doi: <http://www.nature.com/nm/journal/v16/n12/abs/nm.2251.html#supplementary-information>

- Jandrig, B., Seitz, S., Hinzmänn, B., Arnold, W., Micheel, B., Koelble, K., Rosenthal, A. (2004). ST18 is a breast cancer tumor suppressor gene at human chromosome 8q11.2. *Oncogene*, 23(57), 9295-9302.
- Jansen, J. C., Turner, J., Sheehy, J., & Fagan, P. A. (2003). Recurrent Rhabdoid Meningioma: Case Report. *Skull Base*, 13(1), 51-54. doi: 10.1055/s-2003-820557
- Jeibmann, A., Eikmeier, K., Linge, A., Kool, M., Koos, B., Schulz, J., Hasselblatt, M. (2014). Identification of genes involved in the biology of atypical teratoid/rhabdoid tumours using *Drosophila melanogaster*. *Nat Commun*, 5, 4005. doi: 10.1038/ncomms5005
- Jeibmann, A., Eikmeier, K., Linge, A., Kool, M., Koos, B., Schulz, J., Hasselblatt, M. (2014). Identification of genes involved in the biology of atypical teratoid/rhabdoid tumours using *Drosophila melanogaster*. *Nat Commun*, 5. doi: 10.1038/ncomms5005
- Jeong, S. M., Lee, C., Lee, S. K., Kim, J., & Seong, R. H. (2010). The SWI/SNF chromatin-remodeling complex modulates peripheral T cell activation and proliferation by controlling AP-1 expression. *J Biol Chem*, 285(4), 2340-2350. doi: 10.1074/jbc.M109.026997
- Jia, L. Q., Osada, M., Ishioka, C., Gamo, M., Ikawa, S., Suzuki, T., Kuroki, T. (1997). Screening the p53 status of human cell lines using a yeast functional assay. *Mol Carcinog*, 19(4), 243-253.
- Jia, W., Qiu, K., He, M., Song, P., Zhou, Q., Zhou, F., Guo, G. (2013). SOAPfuse: an algorithm for identifying fusion transcripts from paired-end RNA-Seq data. *Genome Biol*, 14(2), R12. doi: 10.1186/gb-2013-14-2-r12
- Johnson, D. S., Mortazavi, A., Myers, R. M., & Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science*, 316(5830), 1497-1502. doi: 10.1126/science.1141319

- Jones, P. (2013). The cancer epigenome. *Genome*, 56(10), 540-541. doi: 10.1139/gen-2013-0147
- Jones, P. A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*, 13(7), 484-492. doi: 10.1038/nrg3230
- Jones, P. A., & Baylin, S. B. (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet*, 3(6), 415-428. doi: 10.1038/nrg816
- Jones, P. A., & Baylin, S. B. (2007). The epigenomics of cancer. *Cell*, 128(4), 683-692. doi: 10.1016/j.cell.2007.01.029
- Judkins, A. R., Burger, P. C., Hamilton, R. L., Kleinschmidt-DeMasters, B., Perry, A., Pomeroy, S. L., Biegel, J. A. (2005). INI1 protein expression distinguishes atypical teratoid/rhabdoid tumor from choroid plexus carcinoma. *J Neuropathol Exp Neurol*, 64(5), 391-397.
- Judkins, A. R., Mauger, J., Ht, A., Rorke, L. B., & Biegel, J. A. (2004). Immunohistochemical analysis of hSNF5/INI1 in pediatric CNS neoplasms. *Am J Surg Pathol*, 28(5), 644-650.
- Kadam, S., & Emerson, B. M. (2003). Transcriptional Specificity of Human SWI/SNF BRG1 and BRM Chromatin Remodeling Complexes. *Mol Cell*, 11(2), 377-389. doi: [http://dx.doi.org/10.1016/S1097-2765\(03\)00034-0](http://dx.doi.org/10.1016/S1097-2765(03)00034-0)
- Kadam, S., McAlpine, G. S., Phelan, M. L., Kingston, R. E., Jones, K. A., & Emerson, B. M. (2000). Functional selectivity of recombinant mammalian SWI/SNF subunits. *Genes Dev*, 14(19), 2441-2451.
- Kadoch, C., Hargreaves, D. C., Hodges, C., Elias, L., Ho, L., Ranish, J., & Crabtree, G. R. (2013). Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet*, 45(6), 592-601. doi: 10.1038/ng.2628

- Kaeser, M. D., Aslanian, A., Dong, M. Q., Yates, J. R., & Emerson, B. M. (2008). BRD7, a novel PBAF-specific SWI/SNF subunit, is required for target gene activation and repression in embryonic stem cells. *J. Biol. Chem.*, 283, 32254-32263.
- Kahali, B., Yu, J., Marquez, S. B., Thompson, K. W., Liang, S. Y., Lu, L., & Reisman, D. (2014). The silencing of the SWI/SNF subunit and anticancer gene BRM in Rhabdoid tumors. *Oncotarget*, 5(10), 3316-3332.
- Kais, Z., Barsky, S. H., Mathsyaraja, H., Zha, A., Ransburgh, D. J., He, G., Parvin, J. D. (2011). KIAA0101 interacts with BRCA1 and regulates centrosome number. *Mol Cancer Res*, 9(8), 1091-1099. doi: 10.1158/1541-7786.mcr-10-0503
- Kang, H., Cui, K., & Zhao, K. (2004). BRG1 controls the activity of the retinoblastoma protein via regulation of p21CIP1/WAF1/SDI. *Mol Cell Biol*, 24(3), 1188-1199.
- Kang, H. J., Park, J. H., Chen, W., Kang, S. I., Moroz, K., Ladanyi, M., & Lee, S. B. (2014). EWS-WT1 oncoprotein activates neuronal reprogramming factor ASCL1 and promotes neural differentiation. *Cancer Res*, 74(16), 4526-4535. doi: 10.1158/0008-5472.can-13-3663
- Kao, C. L., Chiou, S. H., Ho, D. M., Chen, Y. J., Liu, R. S., Lo, C. W., Wong, T. T. (2005). Elevation of plasma and cerebrospinal fluid osteopontin levels in patients with atypical teratoid/rhabdoid tumor. *Am J Clin Pathol*, 123(2), 297-304.
- Karlsson, A., Jonsson, M., Lauss, M., Brunnstrom, H., Jonsson, P., Borg, A., Staaf, J. (2014). Genome-wide DNA methylation analysis of lung carcinoma reveals one neuroendocrine and four adenocarcinoma epitypes associated with patient outcome. *Clin Cancer Res*. doi: 10.1158/1078-0432.ccr-14-1087
- Kasinski, A. L., & Slack, F. J. (2011). Epigenetics and genetics. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy. *Nat Rev Cancer*, 11(12), 849-864. doi: 10.1038/nrc3166

- Kato, H., Honma, R., Sanda, T., Fujiwara, T., Ito, E., Yanagisawa, Y., Watanabe, S. (2007). Knock down of hSNF5/Ini1 causes cell cycle arrest and apoptosis in a p53-dependent manner. *Biochem Biophys Res Commun*, 361(3), 580-585. doi: 10.1016/j.bbrc.2007.07.035
- Katoh, H., Yamashita, K., Waraya, M., Margalit, O., Ooki, A., Tamaki, H., Watanabe, M. (2012). Epigenetic silencing of HOPX promotes cancer progression in colorectal cancer. *Neoplasia*, 14(7), 559-571.
- Katsumi, Y., Iehara, T., Miyachi, M., Yagy, S., Tsubai-Shimizu, S., Kikuchi, K., Hosoi, H. (2011). Sensitivity of malignant rhabdoid tumor cell lines to PD 0332991 is inversely correlated with p16 expression. *Biochem Biophys Res Commun*, 413(1), 62-68. doi: 10.1016/j.bbrc.2011.08.047
- Kenneth, N. S., Mudie, S., van Uden, P., & Rocha, S. (2009). SWI/SNF regulates the cellular response to hypoxia. *J Biol Chem*, 284(7), 4123-4131. doi: 10.1074/jbc.M808491200
- Kia, S. K., Gorski, M. M., Giannakopoulos, S., & Verrijzer, C. P. (2008). SWI/SNF mediates polycomb eviction and epigenetic reprogramming of the INK4b-ARF-INK4a locus. *Mol Cell Biol*, 28(10), 3457-3464. doi: 10.1128/mcb.02019-07
- Kidder, B. L., Hu, G., & Zhao, K. (2011). ChIP-Seq: technical considerations for obtaining high-quality data. *Nat Immunol*, 12(10), 918-922.
- Kidder, B. L., Palmer, S., & Knott, J. G. (2009). SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells. *Stem Cells*, 27(2), 317-328. doi: 10.1634/stemcells.2008-0710
- Kieran, M. W., Roberts, C. W., Chi, S. N., Ligon, K. L., Rich, B. E., Macconail, L. E., Biegel, J. A. (2012). Absence of oncogenic canonical pathway mutations in aggressive pediatric rhabdoid tumors. *Pediatr Blood Cancer*, 59(7), 1155-1157. doi: 10.1002/pbc.24315



- Kim, H., & Park, H. (2007). Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. *Bioinformatics*, 23(12), 1495-1502. doi: 10.1093/bioinformatics/btm134
- Kim, K. H., & Roberts, C. W. (2014). Mechanisms by which SMARCB1 loss drives rhabdoid tumor growth. *Cancer Genet*. doi: 10.1016/j.cancergen.2014.04.004
- Kim, Y., Fedoriw, A. M., & Magnuson, T. (2012). An essential role for a mammalian SWI/SNF chromatin-remodeling complex during male meiosis. *Development*, 139(6), 1133-1140. doi: 10.1242/dev.073478
- Kingston, R. E., & Narlikar, G. J. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev*, 13(18), 2339-2352.
- Klajic, J., Busato, F., Edvardsen, H., Touleimat, N., Fleischer, T., Bukholm, I. R., Kristensen, V. N. (2014). DNA Methylation status of key cell cycle regulators such as CDKN2A/p16 and CCNA1 correlates with treatment response to doxorubicin and 5-fluorouracil in locally advanced breast tumors. *Clin Cancer Res*. doi: 10.1158/1078-0432.ccr-14-0297
- Kleihues, P., Louis, D. N., Scheithauer, B. W., Rorke, L. B., Reifenberger, G., Burger, P. C., & Cavenee, W. K. (2002). The WHO classification of tumors of the nervous system. *J Neuropathol Exp Neurol*, 61(3), 215-225; discussion 226-219.
- Klochender-Yeivin, A., Fiette, L., Barra, J., Muchardt, C., Babinet, C., & Yaniv, M. (2000). The murine SNF5/INI1 chromatin remodeling factor is essential for embryonic development and tumor suppression. *EMBO Rep*, 1(6), 500-506. doi: 10.1093/embo-reports/kvd129
- Klochender-Yeivin, A., Muchardt, C., & Yaniv, M. (2002). SWI/SNF chromatin remodeling and cancer. *Curr Opin Genet Dev*, 12(1), 73-79.

- Klochender-Yeivin, A., Picarsky, E., & Yaniv, M. (2006). Increased DNA damage sensitivity and apoptosis in cells lacking the Snf5/Ini1 subunit of the SWI/SNF chromatin remodeling complex. *Mol Cell Biol*, 26(7), 2661-2674. doi: 10.1128/mcb.26.7.2661-2674.2006
- Knipstein, J. A., Birks, D. K., Donson, A. M., Alimova, I., Foreman, N. K., & Vibhakar, R. (2012). Histone deacetylase inhibition decreases proliferation and potentiates the effect of ionizing radiation in atypical teratoid/rhabdoid tumor cells. *Neuro Oncol*, 14(2), 175-183. doi: 10.1093/neuonc/nor208
- Knipstein, J. A., Birks, D. K., Donson, A. M., Alimova, I., Foreman, N. K., & Vibhakar, R. (2012). Histone deacetylase inhibition decreases proliferation and potentiates the effect of ionizing radiation in atypical teratoid/rhabdoid tumor cells. *Neuro Oncol*, 14(2), 175-183. doi: 10.1093/neuonc/nor208
- Knutson, S. K., Warholic, N. M., Wigle, T. J., Klaus, C. R., Allain, C. J., Raimondi, A., Keilhack, H. (2013). Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. *Proc Natl Acad Sci U S A*, 110(19), 7922-7927. doi: 10.1073/pnas.1303800110
- Ko, F. C., Chan, L. K., Sze, K. M., Yeung, Y. S., Tse, E. Y., Lu, P., Yam, J. W. (2013). PKA-induced dimerization of the RhoGAP DLC1 promotes its inhibition of tumorigenesis and metastasis. *Nat Commun*, 4, 1618. doi: 10.1038/ncomms2604
- Kohashi, K., Nakatsura, T., Kinoshita, Y., Yamamoto, H., Yamada, Y., Tajiri, T., Oda, Y. (2013). Glypican 3 expression in tumors with loss of SMARCB1/INI1 protein expression. *Hum Pathol*, 44(4), 526-533. doi: 10.1016/j.humpath.2012.06.014
- Kordes, U., Bartelheim, K., Modena, P., Massimino, M., Biassoni, V., Reinhard, H., Fruhwald, M. C. (2014). Favorable outcome of patients affected by rhabdoid tumors due to rhabdoid tumor predisposition syndrome (RTPS). *Pediatr Blood Cancer*, 61(5), 919-921. doi: 10.1002/pbc.24793

Kordes, U., Gesk, S., Frühwald, M. C., Graf, N., Leuschner, I., Hasselblatt, M., Schneppenheim, R. (2010). Clinical and molecular features in patients with atypical teratoid rhabdoid tumor or malignant rhabdoid tumor. *Genes, Chromosomes and Cancer*, 49(2), 176-181. doi: 10.1002/gcc.20729

Kosho, T., Miyake, N., & Carey, J. C. (2014). Coffin-Siris syndrome and related disorders involving components of the BAF (mSWI/SNF) complex: Historical review and recent advances using next generation sequencing. *Am J Med Genet C Semin Med Genet*, 166(3), 241-251. doi: 10.1002/ajmg.c.31415

Kosho, T., & Okamoto, N. (2014). Genotype-phenotype correlation of Coffin-Siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A. *Am J Med Genet C Semin Med Genet*, 166(3), 262-275. doi: 10.1002/ajmg.c.31407

Kosho, T., Okamoto, N., Ohashi, H., Tsurusaki, Y., Imai, Y., Hibi-Ko, Y., Matsumoto, N. (2013). Clinical correlations of mutations affecting six components of the SWI/SNF complex: detailed description of 21 patients and a review of the literature. *Am J Med Genet A*, 161a(6), 1221-1237. doi: 10.1002/ajmg.a.35933

Kothandapani, A., Gopalakrishnan, K., Kahali, B., Reisman, D., & Patrick, S. M. (2012). Downregulation of SWI/SNF chromatin remodeling factor subunits modulates cisplatin cytotoxicity. *Exp Cell Res*, 318(16), 1973-1986. doi: 10.1016/j.yexcr.2012.06.011

Kourti, M., Hatzipantelis, E., Zamboukas, T., Tragiannidis, A., Petrakis, G., & Athanassiadou-Piperopoulou, F. (2012). Rare non-Wilms' tumors in children. *Rare Tumors*, 4(1), e6. doi: 10.4081/rt.2012.e6

Kraus, J. A., de Millas, W., Sorensen, N., Herbold, C., Schichor, C., Tonn, J. C., Pietsch, T. (2001). Indications for a tumor suppressor gene at 22q11 involved in the pathogenesis of ependymal tumors and distinct from hSNF5/INI1. *Acta Neuropathol*, 102(1), 69-74.

- Kreiger, P. A., Judkins, A. R., Russo, P. A., Biegel, J. A., Lestini, B. J., Assanasen, C., & Pawel, B. R. (2009). Loss of INI1 expression defines a unique subset of pediatric undifferentiated soft tissue sarcomas. *Mod Pathol*, 22(1), 142-150. doi: 10.1038/modpathol.2008.185
- Ku, C. S., Cooper, D. N., Ziogas, D. E., Halkia, E., Tzaphlidou, M., & Roukos, D. H. (2013). Research and clinical applications of cancer genome sequencing. *Curr Opin Obstet Gynecol*, 25(1), 3-10. doi: 10.1097/GCO.0b013e32835af17c
- Kufe, D. W., Pollock, R. E., & Weichselbaum, R. R. (2003). *Cancer Medicine*. 6th edition (2009/03/24 ed.): BD Decker Ink.
- Kuroda, H., Moritake, H., Sawada, K., Kuwahara, Y., Imoto, I., Inazawa, J., & Sugimoto, T. (2005). Establishment of a cell line from a malignant rhabdoid tumor of the liver lacking the function of two tumor suppressor genes, hSNF5/INI1 and p16. *Cancer Genet Cytogenet*, 158(2), 172-179. doi: 10.1016/j.cancergencyto.2004.08.032
- Kusafuka, T., Miao, J., Yoneda, A., Kuroda, S., & Fukuzawa, M. (2004). Novel germ-line deletion of SNF5/INI1/SMARCB1 gene in neonate presenting with congenital malignant rhabdoid tumor of kidney and brain primitive neuroectodermal tumor. *Genes Chromosomes Cancer*, 40(2), 133-139. doi: 10.1002/gcc.20026
- Kuwahara, Y., Charboneau, A., Knudsen, E. S., & Weissman, B. E. (2010). Reexpression of hSNF5 in malignant rhabdoid tumor cell lines causes cell cycle arrest through a p21(CIP1/WAF1)-dependent mechanism. *Cancer Res*, 70(5), 1854-1865. doi: 10.1158/0008-5472.can-09-1922
- Kuwahara, Y., Mora-Blanco, E. L., Banine, F., Rogers, A. B., Fletcher, C., Sherman, L. S., Weissman, B. E. (2013). Establishment and characterization of MRT cell lines from genetically engineered mouse models and the influence of genetic background on their development. *Int J Cancer*, 132(12), 2767-2777. doi: 10.1002/ijc.27976

- Kuwahara, Y., Wei, D., Durand, J., & Weissman, B. E. (2013). SNF5 reexpression in malignant rhabdoid tumors regulates transcription of target genes by recruitment of SWI/SNF complexes and RNAPII to the transcription start site of their promoters. *Mol Cancer Res*, 11(3), 251-260. doi: 10.1158/1541-7786.mcr-12-0390
- Lan, S. Y., Yu, T., Xia, Z. S., Yuan, Y. H., Shi, L., Lin, Y., Chen, Q. K. (2010). Musashi 1-positive cells derived from mouse embryonic stem cells can differentiate into neural and intestinal epithelial-like cells in vivo. *Cell Biol Int*, 34(12), 1171-1180. doi: 10.1042/cbi20100108
- Lange, C. P., Campan, M., Hinoue, T., Schmitz, R. F., van der Meulen-de Jong, A. E., Slingerland, H., Laird, P. W. (2012). Genome-scale discovery of DNA-methylation biomarkers for blood-based detection of colorectal cancer. *PLoS One*, 7(11), e50266. doi: 10.1371/journal.pone.0050266
- Las Heras, F., & Pritzker, K. P. (2010). Adult variant of atypical teratoid/rhabdoid tumor: immunohistochemical and ultrastructural confirmation of a rare tumor in the sella tursica. *Pathol Res Pract*, 206(11), 788-791. doi: 10.1016/j.prp.2010.07.004
- Le Loarer, F., Zhang, L., Fletcher, C. D., Ribeiro, A., Singer, S., Italiano, A., Antonescu, C. R. (2014). Consistent SMARCB1 homozygous deletions in epithelioid sarcoma and in a subset of myoepithelial carcinomas can be reliably detected by FISH in archival material. *Genes Chromosomes Cancer*, 53(6), 475-486. doi: 10.1002/gcc.22159
- Lee, B., Mukhi, N., & Liu, D. (2012). Current management and novel agents for malignant melanoma. *J Hematol Oncol*, 5, 3. doi: 10.1186/1756-8722-5-3
- Lee, D., Kim, J. W., Seo, T., Hwang, S. G., Choi, E. J., & Choe, J. (2002). SWI/SNF complex interacts with tumor suppressor p53 and is necessary for the activation of p53-mediated transcription. *J Biol Chem*, 277(25), 22330-22337. doi: 10.1074/jbc.M111987200

- Lee, H. Y., Yoon, C. S., Sevenet, N., Rajalingam, V., Delattre, O., & Walford, N. Q. (2002). Rhabdoid tumor of the kidney is a component of the rhabdoid predisposition syndrome. *Pediatr Dev Pathol*, 5(4), 395-399. doi: 10.1007/s10024-001-0259-z
- Lee, J. Y., Kim, I. K., Phi, J. H., Wang, K. C., Cho, B. K., Park, S. H., Kim, S. K. (2012). Atypical teratoid/rhabdoid tumors: the need for more active therapeutic measures in younger patients. *J Neurooncol*, 107(2), 413-419. doi: 10.1007/s11060-011-0769-0
- Lee, K., & Gollahon, L. S. (2014). Zscan4 interacts directly with human Rap1 in cancer cells regardless of telomerase status. *Cancer Biol Ther*, 15(8), 1094-1105. doi: 10.4161/cbt.29220
- Lee, M. C., Park, S. K., Lim, J. S., Jung, S., Kim, J. H., Woo, Y. J., Choi, H. Y. (2002). Atypical teratoid/rhabdoid tumor of the central nervous system: clinico-pathological study. *Neuropathology*, 22(4), 252-260.
- Lee, R. S., & Roberts, C. W. (2013). Rhabdoid tumors: an initial clue to the role of chromatin remodeling in cancer. *Brain Pathol*, 23(2), 200-205. doi: 10.1111/bpa.12021
- Lee, R. S., Stewart, C., Carter, S. L., Ambrogio, L., Cibulskis, K., Sougnez, C., Roberts, C. W. (2012). A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers. *J Clin Invest*, 122(8), 2983-2988. doi: 10.1172/jci64400
- Lee, R. S., Stewart, C., Carter, S. L., Ambrogio, L., Cibulskis, K., Sougnez, C., Roberts, C. W. (2012). A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers. *J Clin Invest*, 122(8), 2983-2988. doi: 10.1172/jci64400
- Lee, S., Cimica, V., Ramachandra, N., Zagzag, D., & Kalpana, G. V. (2011). Aurora A is a repressed effector target of the chromatin remodeling protein INI1/hSNF5 required for rhabdoid tumor cell survival. *Cancer Res*, 71(9), 3225-3235. doi: 10.1158/0008-5472.can-10-2167

- Lei, I., Gao, X., Sham, M. H., & Wang, Z. (2012). SWI/SNF protein component BAF250a regulates cardiac progenitor cell differentiation by modulating chromatin accessibility during second heart field development. *J Biol Chem*, 287(29), 24255-24262. doi: 10.1074/jbc.M112.365080
- Lessard, J. (2007). An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron*, 55, 201-215.
- Levy, P., Ripoché, H., Laurendeau, I., Lazar, V., Ortonne, N., Parfait, B., Bieche, I. (2007). Microarray-based identification of tenascin C and tenascin XB, genes possibly involved in tumorigenesis associated with neurofibromatosis type 1. *Clin Cancer Res*, 13(2 Pt 1), 398-407. doi: 10.1158/1078-0432.ccr-06-0182
- Li, L., Fan, X. S., Xia, Q. Y., Rao, Q., Liu, B., Yu, B., Zhou, X. J. (2014). Concurrent loss of INI1, PBRM1, and BRM expression in epithelioid sarcoma: implications for the cocontributions of multiple SWI/SNF complex members to pathogenesis. *Hum Pathol*. doi: 10.1016/j.humpath.2014.06.027
- Li, Y., Li, S., Chen, J., Shao, T., Jiang, C., Wang, Y., Li, X. (2014). Comparative epigenetic analyses reveal distinct patterns of oncogenic pathways activation in breast cancer subtypes. *Hum Mol Genet*, 23(20), 5378-5393. doi: 10.1093/hmg/ddu256
- Lin, H., Wong, R. P. C., Martinka, M., & Li, G. (2009). Loss of SNF5 expression correlates with poor patient survival in melanoma. *Clinical Cancer Research*, 15(20), 6404-6411.
- Link, A., Balaguer, F., Shen, Y., Lozano, J. J., Leung, H. C., Boland, C. R., & Goel, A. (2013). Curcumin modulates DNA methylation in colorectal cancer cells. *PLoS One*, 8(2), e57709. doi: 10.1371/journal.pone.0057709
- Liu, G., Gramling, S., Munoz, D., Cheng, D., Azad, A. K., Mirshams, M., Reisman, D. (2011). Two novel BRM insertion promoter sequence variants are associated with loss of BRM expression and lung cancer risk. *Oncogene*, 30(29), 3295-3304. doi: 10.1038/onc.2011.81

- Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Law, M. (2012). Comparison of Next-Generation Sequencing Systems. *Journal of Biomedicine and Biotechnology*, 2012, 11. doi: 10.1155/2012/251364
- Lou, J., Wang, Y., Yao, C., Jin, L., Wang, X., Xiao, Y., Zhang, X. (2013). Role of DNA Methylation in Cell Cycle Arrest Induced by Cr (VI) in Two Cell Lines. *PLoS One*, 8(8).
- Louis, D. N., Ohgaki, H., Wiestler, O. D., Cavenee, W. K., Burger, P. C., Jouvett, A., Kleihues, P. (2007). The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol*, 114(2), 97-109. doi: 10.1007/s00401-007-0243-4
- Lu, H. G., Zhan, W., Yan, L., Qin, R. Y., Yan, Y. P., Yang, Z. J., Chen, G. Q. (2014). TET1 partially mediates HDAC inhibitor-induced suppression of breast cancer invasion. *Molecular Medicine Reports*, 10(5), 2595-2600.
- Lu, P., & Roberts, C. W. (2013). The SWI/SNF tumor suppressor complex: Regulation of promoter nucleosomes and beyond. *Nucleus*, 4(5), 374-378. doi: 10.4161/nucl.26654
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8[thinsp]Å resolution. *Nature*, 389(6648), 251-260.
- Ma, H. I., Kao, C. L., Lee, Y. Y., Chiou, G. Y., Tai, L. K., Lu, K. H., Wong, T. T. (2010). Differential expression profiling between atypical teratoid/rhabdoid and medulloblastoma tumor in vitro and in vivo using microarray analysis. *Childs Nerv Syst*, 26(3), 293-303. doi: 10.1007/s00381-009-1016-2
- Ma, H.-I., Kao, C.-L., Lee, Y.-Y., Chiou, G.-Y., Tai, L.-K., Lu, K.-H., Wong, T.-T. (2010). Differential expression profiling between atypical teratoid/rhabdoid and medulloblastoma tumor in vitro and in vivo using microarray analysis. *Child's Nervous System*, 26(3), 293-303. doi: 10.1007/s00381-009-1016-2



- Ma, P., Yang, X., Kong, Q., Li, C., Yang, S., Li, Y., & Mao, B. (2014). The Ubiquitin Ligase RNF220 Enhances Canonical Wnt Signaling through USP7-Mediated Deubiquitination of beta-Catenin. *Mol Cell Biol*, 34(23), 4355-4366. doi: 10.1128/mcb.00731-14
- MacDonald, T. J. (2008). Aggressive infantile embryonal tumors. *J Child Neurol*, 23(10), 1195-1204. doi: 10.1177/0883073808321769
- Macurek, L., Lindqvist, A., Lim, D., Lampson, M. A., Klompmaker, R., Freire, R., Medema, R. H. (2008). Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery. *Nature*, 455(7209), 119-123. doi: 10.1038/nature07185
- Maher, C. A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., Chinnaiyan, A. M. (2009). Transcriptome sequencing to detect gene fusions in cancer. *Nature*, 458(7234), 97-101. doi: 10.1038/nature07638
- Majidi, M., Hubbs, A. E., & Lichy, J. H. (1998). Activation of extracellular signal-regulated kinase 2 by a novel Abl-binding protein, ST5. *J Biol Chem*, 273(26), 16608-16614.
- Majidi, M., Hubbs, A. E., & Lichy, J. H. (1998). Activation of extracellular signal-regulated kinase 2 by a novel Abl-binding protein, ST5. *J Biol Chem*, 273(26), 16608-16614.
- Maksimovic, J., Gordon, L., & Oshlack, A. (2012). SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol*, 13(6), R44. doi: 10.1186/gb-2012-13-6-r44
- Martens, J. A., & Winston, F. (2002). Evidence that Swi/Snf directly represses transcription in *S. cerevisiae*. *Genes Dev*, 16(17), 2231-2236. doi: 10.1101/gad.1009902
- Martens, J. A., & Winston, F. (2003). Recent advances in understanding chromatin remodeling by Swi/Snf complexes. *Curr Opin Genet Dev*, 13(2), 136-142.

- Martin, J. A., & Wang, Z. (2011). Next-generation transcriptome assembly. *Nat Rev Genet*, 12(10), 671-682.
- Martin-Subero, J. I., Ammerpohl, O., Bibikova, M., Wickham-Garcia, E., Agirre, X., Alvarez, S., Siebert, R. (2009). A comprehensive microarray-based DNA methylation study of 367 hematological neoplasms. *PLoS One*, 4(9), e6986. doi: 10.1371/journal.pone.0006986
- Mbandi, S. K., Hesse, U., Rees, D. J., & Christoffels, A. (2014). A glance at quality score: implication for de novo transcriptome reconstruction of Illumina reads. *Front Genet*, 5, 17. doi: 10.3389/fgene.2014.00017
- McIlwain, D. R., Berger, T., & Mak, T. W. (2013). Caspase functions in cell death and disease. *Cold Spring Harb Perspect Biol*, 5(4), a008656. doi: 10.1101/cshperspect.a008656
- McKenna, E. S., Sansam, C. G., Cho, Y. J., Greulich, H., Evans, J. A., Thom, C. S., Roberts, C. W. (2008). Loss of the epigenetic tumor suppressor SNF5 leads to cancer without genomic instability. *Mol Cell Biol*, 28(20), 6223-6233. doi: 10.1128/mcb.00658-08
- McKenna, E. S., Tamayo, P., Cho, Y. J., Tillman, E. J., Mora-Blanco, E. L., Sansam, C. G., Roberts, C. W. (2012). Epigenetic inactivation of the tumor suppressor BIN1 drives proliferation of SNF5-deficient tumors. *Cell Cycle*, 11(10), 1956-1965. doi: 10.4161/cc.20280
- McLendon, R. E., Adekunle, A., Rajaram, V., Kocak, M., & Blaney, S. M. (2011). Embryonal central nervous system neoplasms arising in infants and young children: a pediatric brain tumor consortium study. *Arch Pathol Lab Med*, 135(8), 984-993. doi: 10.5858/2010-0515-oar1
- McShane, L. M., Radmacher, M. D., Freidlin, B., Yu, R., Li, M. C., & Simon, R. (2002). Methods for assessing reproducibility of clustering patterns observed in analyses of microarray data. *Bioinformatics*, 18(11), 1462-1469.

- Mechlin, C. W., Tanner, M. J., Chen, M., Buttyan, R., Levin, R. M., & Mian, B. M. (2010). Gli2 expression and human bladder transitional carcinoma cell invasiveness. *J Urol*, 184(1), 344-351. doi: 10.1016/j.juro.2010.03.007
- Medina, P. P., Carretero, J., Fraga, M. F., Esteller, M., Sidransky, D., & Sanchez-Cespedes, M. (2004). Genetic and epigenetic screening for gene alterations of the chromatin-remodeling factor, SMARCA4/BRG1, in lung tumors. *Genes Chromosomes Cancer*, 41(2), 170-177. doi: 10.1002/gcc.20068
- Medina, P. P., Romero, O. A., Kohno, T., Montuenga, L. M., Pio, R., Yokota, J., & Sanchez-Cespedes, M. (2008). Frequent BRG1/SMARCA4-inactivating mutations in human lung cancer cell lines. *Hum Mutat*, 29(5), 617-622. doi: 10.1002/humu.20730
- Medina, P. P., & Sanchez-Cespedes, M. (2008). Involvement of the chromatin-remodeling factor BRG1/SMARCA4 in human cancer. *Epigenetics*, 3(2), 64-68.
- Medjkane, S., Novikov, E., Versteeg, I., & Delattre, O. (2004). The tumor suppressor hSNF5/INI1 modulates cell growth and actin cytoskeleton organization. *Cancer Res*, 64(10), 3406-3413. doi: 10.1158/0008-5472.can-03-3004
- Meldrum, C., Doyle, M. A., & Tothill, R. W. (2011). Next-generation sequencing for cancer diagnostics: a practical perspective. *Clin Biochem Rev*, 32(4), 177-195.
- Merhavi, E., Cohen, Y., Avraham, B. C., Frenkel, S., Chowers, I., Pe'er, J., & Goldenberg-Cohen, N. (2007). Promoter methylation status of multiple genes in uveal melanoma. *Invest Ophthalmol Vis Sci*, 48(10), 4403-4406. doi: 10.1167/iovs.07-0272
- Mills, A. A. (2010). Throwing the cancer switch: reciprocal roles of polycomb and trithorax proteins. *Nat Rev Cancer*, 10(10), 669-682.
- Mine, N., Bando, K., Utada, Y., Nagai, H., Araki, T., & Emi, M. (1999). Two single nucleotide polymorphisms of the hSNF5/INI1 gene. *J Hum Genet*, 44(5), 354-355. doi: 10.1007/s100380050177

- Miranda, T. B., & Jones, P. A. (2007). DNA methylation: the nuts and bolts of repression. *J Cell Physiol*, 213(2), 384-390. doi: 10.1002/jcp.21224
- Miremadi, A., Oestergaard, M. Z., Pharoah, P. D. P., & Caldas, C. (2007). Cancer genetics of epigenetic genes. *Hum Mol Genet*, 16(R1), R28-R49. doi: 10.1093/hmg/ddm021
- Mitrofanova, A. M., Konovalov, D. M., Kisliakov, A. N., Roshchin, V., & Abramov, D. S. (2013). [Atypical teratoid/rhabdoid tumors of childhood]. *Arkh Patol*, 75(5), 36-42.
- Miyake, N., Tsurusaki, Y., & Matsumoto, N. (2014). Numerous BAF complex genes are mutated in Coffin-Siris syndrome. *Am J Med Genet C Semin Med Genet*, 166(3), 257-261. doi: 10.1002/ajmg.c.31406
- Mobley, B. C., McKenney, J. K., Bangs, C. D., Callahan, K., Yeom, K. W., Schneppenheim, R., Vogel, H. (2010). Loss of SMARCB1/INI1 expression in poorly differentiated chordomas. *Acta Neuropathol*, 120(6), 745-753. doi: 10.1007/s00401-010-0767-x
- Modena, P., Lualdi, E., Facchinetti, F., Galli, L., Teixeira, M. R., Pilotti, S., & Sozzi, G. (2005). SMARCB1/INI1 tumor suppressor gene is frequently inactivated in epithelioid sarcomas. *Cancer Res*, 65(10), 4012-4019. doi: 10.1158/0008-5472.can-04-3050
- Modena, P., Lualdi, E., Facchinetti, F., Galli, L., Teixeira, M. R., Pilotti, S., & Sozzi, G. (2005). SMARCB1/INI1 tumor suppressor gene is frequently inactivated in epithelioid sarcomas. *Cancer Res*, 65(10), 4012-4019. doi: 10.1158/0008-5472.can-04-3050
- Modena, P., Sardi, I., Brenca, M., Giunti, L., Buccoliero, A. M., Pollo, B., Massimino, M. (2013). Case report: long-term survival of an infant syndromic patient affected by atypical teratoid-rhabdoid tumor. *BMC Cancer*, 13, 100. doi: 10.1186/1471-2407-13-100

- Mokry, M., Hatzis, P., Schuijers, J., Lansu, N., Ruzius, F. P., Clevers, H., & Cuppen, E. (2012). Integrated genome-wide analysis of transcription factor occupancy, RNA polymerase II binding and steady-state RNA levels identify differentially regulated functional gene classes. *Nucleic Acids Res*, 40(1), 148-158. doi: 10.1093/nar/gkr720
- Moretti, C., Lupoi, D., Spasaro, F., Chioma, L., Di Giacinto, P., Colicchia, M., Gnessi, L. (2013). Sella turcica atypical teratoid/rhabdoid tumor complicated with lung metastasis in an adult female. *Clin Med Insights Case Rep*, 6, 177-182. doi: 10.4137/CCRep.S12834
- Morgenstern, D. A., Gibson, S., Brown, T., Sebire, N. J., & Anderson, J. (2010). Clinical and pathological features of paediatric malignant rhabdoid tumours. *Pediatr Blood Cancer*, 54(1), 29-34. doi: 10.1002/pbc.22231
- Morozov, A., Lee, S. J., Zhang, Z. K., Cimica, V., Zagzag, D., & Kalpana, G. V. (2007). INI1 induces interferon signaling and spindle checkpoint in rhabdoid tumors. *Clin Cancer Res*, 13(16), 4721-4730. doi: 10.1158/1078-0432.ccr-07-0054
- Morris, T. J., & Beck, S. Analysis pipelines and packages for Infinium HumanMethylation450 BeadChip (450k) data. *Methods*(0). doi: <http://dx.doi.org/10.1016/j.ymeth.2014.08.011>
- Muchardt, C., & Yaniv, M. (1999). The mammalian SWI/SNF complex and the control of cell growth. *Semin Cell Dev Biol*, 10(2), 189-195. doi: 10.1006/scdb.1999.0300
- Muchardt, C., & Yaniv, M. (2001). When the SWI/SNF complex remodels...the cell cycle. *Oncogene*, 20(24), 3067-3075. doi: 10.1038/sj.onc.1204331
- Muhlisch, J., Schwering, A., Grotzer, M., Vince, G. H., Roggendorf, W., Hagemann, C., Fruhwald, M. C. (2005). Epigenetic repression of RASSF1A but not CASP8 in supratentorial PNET (sPNET) and atypical teratoid/rhabdoid tumors (AT/RT) of childhood. *Oncogene*, 25(7), 1111-1117.

- Muhlisch, J., Schwering, A., Grotzer, M., Vince, G. H., Roggendorf, W., Hagemann, C., Fruhwald, M. C. (2005). Epigenetic repression of RASSF1A but not CASP8 in supratentorial PNET (sPNET) and atypical teratoid/rhabdoid tumors (AT/RT) of childhood. *Oncogene*, 25(7), 1111-1117.
- Muller, J., & Verrijzer, P. (2009). Biochemical mechanisms of gene regulation by polycomb group protein complexes. *Curr Opin Genet Dev*, 19(2), 150-158. doi: 10.1016/j.gde.2009.03.001
- Murthi, P. (2014). Review: placental homeobox genes and their role in regulating human fetal growth. *Placenta*, 35 Suppl, S46-50. doi: 10.1016/j.placenta.2013.11.006
- Narendran, A., Coppes, L., Jayanthan, A., Coppes, M., Teja, B., Bernoux, D., Strother, D. (2008). Establishment of atypical-teratoid/rhabdoid tumor (AT/RT) cell cultures from disseminated CSF cells: a model to elucidate biology and potential targeted therapeutics. *J Neurooncol*, 90(2), 171-180. doi: 10.1007/s11060-008-9653-y
- Narlikar, G. J., Fan, H. Y., & Kingston, R. E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. *Cell*, 108(4), 475-487.
- Narlikar, Geeta J., Sundaramoorthy, R., & Owen-Hughes, T. Mechanisms and Functions of ATP-Dependent Chromatin-Remodeling Enzymes. *Cell*, 154(3), 490-503. doi: 10.1016/j.cell.2013.07.011
- Naumov, V. A., Generozov, E. V., Zaharjevskaya, N. B., Matushkina, D. S., Larin, A. K., Chernyshov, S. V., Govorun, V. M. (2013). Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips. *Epigenetics*, 8(9), 921-934. doi: 10.4161/epi.25577
- Ndlovu, M. N., Denis, H., & Fuks, F. (2011). Exposing the DNA methylome iceberg. *Trends in Biochemical Sciences*, 36(7), 381-387. doi: <http://dx.doi.org/10.1016/j.tibs.2011.03.002>

- Ni, Z., & Bremner, R. (2007). Brahma-related gene 1-dependent STAT3 recruitment at IL-6-inducible genes. *J Immunol*, 178(1), 345-351.
- Ni, Z., & Bremner, R. (2007). Brahma-related gene 1-dependent STAT3 recruitment at IL-6-inducible genes. *J Immunol*, 178(1), 345-351.
- Nie, Z., Xue, Y., Yang, D., Zhou, S., Deroo, B. J., Archer, T. K., & Wang, W. (2000). A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex. *Mol Cell Biol*, 20(23), 8879-8888.
- Nie, Z., Yan, Z., Chen, E. H., Sechi, S., Ling, C., Zhou, S., Wang, W. (2003). Novel SWI/SNF chromatin-remodeling complexes contain a mixed-lineage leukemia chromosomal translocation partner. *Mol Cell Biol*, 23(8), 2942-2952.
- Northrup, D. L., & Zhao, K. (2011). Application of ChIP-Seq and related techniques to the study of immune function. *Immunity*, 34(6), 830-842. doi: 10.1016/j.immuni.2011.06.002
- Oh, J., Sohn, D. H., Ko, M., Chung, H., Jeon, S. H., & Seong, R. H. (2008). BAF60a interacts with p53 to recruit the SWI/SNF complex. *J Biol Chem*, 283(18), 11924-11934. doi: 10.1074/jbc.M705401200
- Oike, T., Ogiwara, H., Nakano, T., Yokota, J., & Kohno, T. (2013). Inactivating mutations in SWI/SNF chromatin remodeling genes in human cancer. *Jpn J Clin Oncol*, 43(9), 849-855. doi: 10.1093/jjco/hyt101
- Okabe, H., Lee, S. H., Phuchareon, J., Albertson, D. G., McCormick, F., & Tetsu, O. (2006). A critical role for FBXW8 and MAPK in cyclin D1 degradation and cancer cell proliferation. *PLoS One*, 1, e128. doi: 10.1371/journal.pone.0000128
- Ooki, A., Yamashita, K., Kikuchi, S., Sakuramoto, S., Katada, N., Kokubo, K., Watanabe, M. (2010). Potential utility of HOP homeobox gene promoter methylation as a marker of tumor aggressiveness in gastric cancer. *Oncogene*, 29(22), 3263-3275. doi: 10.1038/onc.2010.76

- Orlando, V., & Paro, R. (1995). Chromatin multiprotein complexes involved in the maintenance of transcription patterns. *Curr Opin Genet Dev*, 5(2), 174-179.
- Oruetebarria, I., Venturini, F., Kekarainen, T., Houweling, A., Zuijderduijn, L. M., Mohd-Sarip, A., Verrijzer, C. P. (2004). P16INK4a is required for hSNF5 chromatin remodeler-induced cellular senescence in malignant rhabdoid tumor cells. *J Biol Chem*, 279(5), 3807-3816. doi: 10.1074/jbc.M309333200
- Oshlack, A., Robinson, M. D., & Young, M. D. (2010). From RNA-seq reads to differential expression results. *Genome Biol*, 11(12), 220. doi: 10.1186/gb-2010-11-12-220
- Ostrom, Q. T., Chen, Y., P, M. d. B., Ondracek, A., Farah, P., Gittleman, H., Barnholtz-Sloan, J. S. (2014). The descriptive epidemiology of atypical teratoid/rhabdoid tumors in the United States, 2001-2010. *Neuro Oncol*, 16(10), 1392-1399. doi: 10.1093/neuonc/nou090
- P, W., N, F., H, S., G, G., & O, H. (2013). The SWI/SNF chromatin remodeling complex selectively affects multiple aspects of serotonergic neuron differentiation. *Genetics*, 194(1), 189-198. doi: 10.1534/genetics.112.148742
- Pacheco-Pinedo, E. C., Durham, A. C., Stewart, K. M., Goss, A. M., Lu, M. M., DeMayo, F. J., & Morrissey, E. E. (2011). Wnt/ $\beta$ -catenin signaling accelerates mouse lung tumorigenesis by imposing an embryonic distal progenitor phenotype on lung epithelium. *J Clin Invest*, 121(5), 1935-1945. doi: 10.1172/JCI44871
- Packer, R. J. (2008). Childhood brain tumors: accomplishments and ongoing challenges. *J Child Neurol*, 23(10), 1122-1127. doi: 10.1177/0883073808320758
- Packer, R. J., Biegel, J. A., Blaney, S., Finlay, J., Geyer, J. R., Heideman, R., Smith, M. (2002). Atypical teratoid/rhabdoid tumor of the central nervous system: report on workshop. *J Pediatr Hematol Oncol*, 24(5), 337-342.



Packer, R. J., Cogen, P., Vezina, G., & Rorke, L. B. (1999). Medulloblastoma: clinical and biologic aspects. *Neuro Oncol*, 1(3), 232-250.

Pai Panandiker, A. S., Merchant, T. E., Beltran, C., Wu, S., Sharma, S., Boop, F. A., Gajjar, A. (2012). Sequencing of local therapy affects the pattern of treatment failure and survival in children with atypical teratoid rhabdoid tumors of the central nervous system. *Int J Radiat Oncol Biol Phys*, 82(5), 1756-1763. doi: 10.1016/j.ijrobp.2011.02.059

Park, E. S., Sung, K. W., Baek, H. J., Park, K. D., Park, H. J., Won, S. C., Kim, H. S. (2012). Tandem high-dose chemotherapy and autologous stem cell transplantation in young children with atypical teratoid/rhabdoid tumor of the central nervous system. *J Korean Med Sci*, 27(2), 135-140. doi: 10.3346/jkms.2012.27.2.135

Park, J. H., Park, E. J., Hur, S. K., Kim, S., & Kwon, J. (2009). Mammalian SWI/SNF chromatin remodeling complexes are required to prevent apoptosis after DNA damage. *DNA Repair (Amst)*, 8(1), 29-39. doi: 10.1016/j.dnarep.2008.08.011

Park, J. Y., Lee, J. E., Park, J. B., Yoo, H., Lee, S. H., & Kim, J. H. (2014). Roles of Long Non-Coding RNAs on Tumorigenesis and Glioma Development. *Brain Tumor Res Treat*, 2(1), 1-6. doi: 10.14791/btrt.2014.2.1.1

Park, P. J. (2009). ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet*, 10(10), 669-680. doi: 10.1038/nrg2641

Parsons, D. W., Li, M., Zhang, X., Jones, S., Leary, R. J., Lin, J. C., Velculescu, V. E. (2011). The genetic landscape of the childhood cancer medulloblastoma. *Science*, 331(6016), 435-439. doi: 10.1126/science.1198056

Parwani, A. V., Stelow, E. B., Pambuccian, S. E., Burger, P. C., & Ali, S. Z. (2005). Atypical teratoid/rhabdoid tumor of the brain: cytopathologic characteristics and differential diagnosis. *Cancer*, 105(2), 65-70. doi: 10.1002/cncr.20872

- Peris-Bonet, R., Martinez-Garcia, C., Lacour, B., Petrovich, S., Giner-Ripoll, B., Navajas, A., & Steliarova-Foucher, E. (2006). Childhood central nervous system tumours--incidence and survival in Europe (1978-1997): report from Automated Childhood Cancer Information System project. *Eur J Cancer*, 42(13), 2064-2080. doi: 10.1016/j.ejca.2006.05.009
- Pfister, A. S., Tanneberger, K., Schambony, A., & Behrens, J. (2012). Amer2 protein is a novel negative regulator of Wnt/beta-catenin signaling involved in neuroectodermal patterning. *J Biol Chem*, 287(3), 1734-1741. doi: 10.1074/jbc.M111.308650
- Pfister, S. M., Korshunov, A., Kool, M., Hasselblatt, M., Eberhart, C., & Taylor, M. D. (2010). Molecular diagnostics of CNS embryonal tumors. *Acta Neuropathol*, 120(5), 553-566. doi: 10.1007/s00401-010-0751-5
- Phelan, M. L., Sif, S., Narlikar, G. J., & Kingston, R. E. (1999). Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol Cell*, 3(2), 247-253.
- Polytarchou, C., Koukos, G., & Iliopoulos, D. (2014). Systems biology in inflammatory bowel diseases: Ready for prime time. *Current Opinion in Gastroenterology*, 30(4), 339-346.
- Portela, A., & Esteller, M. (2010). Epigenetic modifications and human disease. *Nat Biotech*, 28(10), 1057-1068. doi: 10.1038/nbt.1685
- Poschl, J., Lorenz, A., Hartmann, W., von Bueren, A. O., Kool, M., Li, S., Schuller, U. (2011). Expression of BARHL1 in medulloblastoma is associated with prolonged survival in mice and humans. *Oncogene*, 30(47), 4721-4730. doi: <http://www.nature.com/onc/journal/v30/n47/supinfo/onc2011173s1.html>
- Pottier, N., Yang, W., Assem, M., Panetta, J. C., Pei, D., Paugh, S. W., Cheok, M. H. (2008). The SWI/SNF chromatin-remodeling complex and glucocorticoid resistance in acute lymphoblastic leukemia. *J Natl Cancer Inst*, 100(24), 1792-1803. doi: 10.1093/jnci/djn416

- Puda, A., Milosevic, J. D., Berg, T., Klampfl, T., Harutyunyan, A. S., Gisslinger, B., Kralovics, R. (2012). Frequent deletions of JARID2 in leukemic transformation of chronic myeloid malignancies. *Am J Hematol*, 87(3), 245-250. doi: 10.1002/ajh.22257
- Qu, D., May, R., Sureban, S. M., Weygant, N., Chandrakesan, P., Ali, N., Houchen, C. W. (2014). Inhibition of Notch signaling reduces the number of surviving Dclk1+ reserve crypt epithelial stem cells following radiation injury. *Am J Physiol Gastrointest Liver Physiol*, 306(5), G404-411. doi: 10.1152/ajpgi.00088.2013
- Quinlan, A. R., & Hall, I. M. (2012). Characterizing complex structural variation in germline and somatic genomes. *Trends Genet*, 28(1), 43-53. doi: 10.1016/j.tig.2011.10.002
- Ramanan, V. K., Shen, L., Moore, J. H., & Saykin, A. J. Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends in Genetics*, 28(7), 323-332. doi: 10.1016/j.tig.2012.03.004
- Ranneberg-Nilsen, T., Rollag, H., Slettebakk, R., Backe, P. H., Olsen, O., Luna, L., & Bjoras, M. (2012). The chromatin remodeling factor SMARCB1 forms a complex with human cytomegalovirus proteins UL114 and UL44. *PLoS One*, 7(3), e34119. doi: 10.1371/journal.pone.0034119
- Rao, Q., Xia, Q. Y., Shen, Q., Shi, S. S., Tu, P., Shi, Q. L., & Zhou, X. J. (2014). Coexistent loss of INI1 and BRG1 expression in a rhabdoid renal cell carcinoma (RCC): implications for a possible role of SWI/SNF complex in the pathogenesis of RCC. *Int J Clin Exp Pathol*, 7(4), 1782-1787.
- Ratovitski, E. A. (2013). Tumor protein p63/microRNA network in epithelial cancer cells. *Current Genomics*, 14(7), 441-452.
- Reddy, A. T. (2005). Atypical teratoid/rhabdoid tumors of the central nervous system. *J Neurooncol*, 75(3), 309-313. doi: 10.1007/s11060-005-6762-8

- Reincke, B. S., Rosson, G. B., Oswald, B. W., & Wright, C. F. (2003). INI1 expression induces cell cycle arrest and markers of senescence in malignant rhabdoid tumor cells. *J Cell Physiol*, 194(3), 303-313. doi: 10.1002/jcp.10201
- Reinhard, H., Reinert, J., Beier, R., Furtwangler, R., Alkasser, M., Rutkowski, S., Graf, N. (2008). Rhabdoid tumors in children: prognostic factors in 70 patients diagnosed in Germany. *Oncol Rep*, 19(3), 819-823.
- Reisman, D., Glaros, S., & Thompson, E. A. (2009). The SWI/SNF complex and cancer. *Oncogene*, 28(14), 1653-1668. doi: 10.1038/onc.2009.4
- Reisman, D. N., Sciarrotta, J., Bouldin, T. W., Weissman, B. E., & Funkhouser, W. K. (2005). The expression of the SWI/SNF ATPase subunits BRG1 and BRM in normal human tissues. *Appl Immunohistochem Mol Morphol*, 13(1), 66-74.
- Reisman, D. N., Sciarrotta, J., Wang, W., Funkhouser, W. K., & Weissman, B. E. (2003). Loss of BRG1/BRM in human lung cancer cell lines and primary lung cancers: correlation with poor prognosis. *Cancer Res*, 63(3), 560-566.
- Renard, C., Pissaloux, D., Decouvelaere, A. V., Bourdeaut, F., & Ranchere, D. (2014). Non-rhabdoid pediatric SMARCB1-deficient tumors: overlap between chordomas and malignant rhabdoid tumors? *Cancer Genet*. doi: 10.1016/j.cancergen.2014.05.005
- Reyes, J. C., Barra, J., Muchardt, C., Camus, A., Babinet, C., & Yaniv, M. (1998). Altered control of cellular proliferation in the absence of mammalian brahma (SNF2alpha). *Embo j*, 17(23), 6979-6991. doi: 10.1093/emboj/17.23.6979
- Rezza, A., Skah, S., Roche, C., Nadjari, J., Samarut, J., & Plateroti, M. (2010). The overexpression of the putative gut stem cell marker Musashi-1 induces tumorigenesis through Wnt and Notch activation. *J Cell Sci*, 123(Pt 19), 3256-3265. doi: 10.1242/jcs.065284

- Riyadh, M. A., Shinmyo, Y., Ohta, K., & Tanaka, H. (2014). Inhibitory effects of draxin on axonal outgrowth and migration of precerebellar neurons. *Biochem Biophys Res Commun*, 449(1), 169-174. doi: 10.1016/j.bbrc.2014.05.013
- Rizzo, D., Freneaux, P., Brisse, H., Louvrier, C., Lequin, D., Nicolas, A., Bourdeaut, F. (2012). SMARCB1 deficiency in tumors from the peripheral nervous system: a link between schwannomas and rhabdoid tumors? *Am J Surg Pathol*, 36(7), 964-972. doi: 10.1097/PAS.0b013e31825798f1
- Roberts, A., Pimentel, H., Trapnell, C., & Pachter, L. (2011). Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics*, 27(17), 2325-2329. doi: 10.1093/bioinformatics/btr355
- Roberts, C. W., & Biegel, J. A. (2009). The role of SMARCB1/INI1 in development of rhabdoid tumor. *Cancer Biol Ther*, 8(5), 412-416.
- Roberts, C. W., Leroux, M. M., Fleming, M. D., & Orkin, S. H. (2002). Highly penetrant, rapid tumorigenesis through conditional inversion of the tumor suppressor gene *Snf5*. *Cancer Cell*, 2(5), 415-425.
- Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Jones, S. (2007). Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Meth*, 4(8), 651-657. doi: [http://www.nature.com/nmeth/journal/v4/n8/supinfo/nmeth1068\\_S1.html](http://www.nature.com/nmeth/journal/v4/n8/supinfo/nmeth1068_S1.html)
- Robinson, M. D., Stirzaker, C., Statham, A. L., Coolen, M. W., Song, J. Z., Nair, S. S., Clark, S. J. (2010). Evaluation of affinity-based genome-wide DNA methylation data: effects of CpG density, amplification bias, and copy number variation. *Genome Res*, 20(12), 1719-1729. doi: 10.1101/gr.110601.110
- Rodriguez-Paredes, M., & Esteller, M. (2011). Cancer epigenetics reaches mainstream oncology. *Nat Med*, 330-339.

Ronald, B. (2005). *Child and adolescent neurology Pathol Res Pract* (2009/02/27 ed., pp. 105). London: Wiley-Blackwell

Ronan, J. L., Wu, W., & Crabtree, G. R. (2013). From neural development to cognition: Unexpected roles for chromatin. *Nature Reviews Genetics*, 14(5), 347-359.

Rorke, L. B., Packer, R. J., & Biegel, J. A. (1996). Central nervous system atypical teratoid/rhabdoid tumors of infancy and childhood: definition of an entity. *J Neurosurg*, 85(1), 56-65. doi: 10.3171/jns.1996.85.1.0056

Ross-Innes, C. S., Stark, R., Teschendorff, A. E., Holmes, K. A., Ali, H. R., Dunning, M. J., Carroll, J. S. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*, 481(7381), 389-393. doi: <http://www.nature.com/nature/journal/v481/n7381/abs/nature10730.html#supplementary-information>

Rota, R., Ciarapica, R., Miele, L., & Locatelli, F. (2012). Notch signaling in pediatric soft tissue sarcomas. *BMC Med*, 10, 141. doi: 10.1186/1741-7015-10-141

Rousseau, G., Noguchi, T., Bourdon, V., Sobol, H., & Olschwang, S. (2011). SMARCB1/INI1 germline mutations contribute to 10% of sporadic schwannomatosis. *BMC Neurol*, 11, 9. doi: 10.1186/1471-2377-11-9

Rousseau-Merck, M. F., Versteeg, I., Legrand, I., Couturier, J., Mairal, A., Delattre, O., & Aurias, A. (1999). hSNF5/INI1 inactivation is mainly associated with homozygous deletions and mitotic recombinations in rhabdoid tumors. *Cancer Res*, 59(13), 3152-3156.

Rousseau-Merck, M. F., Versteeg, I., Legrand, I., Couturier, J., Mairal, A., Delattre, O., & Aurias, A. (1999). hSNF5/INI1 inactivation is mainly associated with homozygous deletions and mitotic recombinations in rhabdoid tumors. *Cancer Res*, 59(13), 3152-3156.

- Ryan, M. P., Jones, R., & Morse, R. H. (1998). SWI-SNF complex participation in transcriptional activation at a step subsequent to activator binding. *Mol Cell Biol*, 18(4), 1774-1782.
- Ryan, M. P., Jones, R., & Morse, R. H. (1998). SWI-SNF complex participation in transcriptional activation at a step subsequent to activator binding. *Mol Cell Biol*, 18(4), 1774-1782.
- Saha, A., Wittmeyer, J., & Cairns, B. R. (2005). Chromatin remodeling through directional DNA translocation from an internal nucleosomal site. *Nat Struct Mol Biol*, 12(9), 747-755. doi: 10.1038/nsmb973
- Saha, A., Wittmeyer, J., & Cairns, B. R. (2006). Chromatin remodelling: The industrial revolution of DNA around histones. *Nature Reviews Molecular Cell Biology*, 7(6), 437-447.
- Saito, Y., & Jones, P. A. (2006). Epigenetic activation of tumor suppressor microRNAs in human cancer cells. *Cell Cycle*, 5(19), 2220-2222.
- Samaras, V., Stamatelli, A., Samaras, E., Stergiou, I., Konstantopoulou, P., Varsos, V., Barbatis, C. (2009). Atypical teratoid/rhabdoid tumor of the central nervous system in an 18-year-old patient. *Clin Neuropathol*, 28(1), 1-10.
- San Jose-Eneriz, E., Agirre, X., Rodríguez-Otero, P., & Prosper, F. (2013). Epigenetic regulation of cell signaling pathways in acute lymphoblastic leukemia. *Epigenomics*, 5(5), 525-538.
- Sandoval, J., Heyn, H., Moran, S., Serra-Musach, J., Pujana, M. A., Bibikova, M., & Esteller, M. (2011). Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics*, 6(6), 692-702.
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 74(12), 5463-5467.

- Sansam, C. G., & Roberts, C. W. (2006). Epigenetics and cancer: altered chromatin remodeling via Snf5 loss leads to aberrant cell cycle regulation. *Cell Cycle*, 5(6), 621-624.
- Santen, G. W., Aten, E., Vulto-van Silfhout, A. T., Pottinger, C., van Bon, B. W., van Minderhout, I. J., van Belzen, M. J. (2013). Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. *Hum Mutat*, 34(11), 1519-1528. doi: 10.1002/humu.22394
- Santen, G. W., Kriek, M., & van Attikum, H. (2012). SWI/SNF complex in disorder: SWItching from malignancies to intellectual disability. *Epigenetics*, 7(11), 1219-1224. doi: 10.4161/epi.22299
- Santos, G. C., Carvalho, K. C., Falzoni, R., Simoes, A. C. Q., Rocha, R. M., Lopes, A., da Cunha, I. W. (2009). Glial fibrillary acidic protein in tumor types with cartilaginous differentiation. *Mod Pathol*, 22(10), 1321-1327.
- Sathyamoorthi, S., Morales, J., Bermudez, J., McBride, L., Luquette, M., McGoey, R., Lacassie, Y. (2009). Array analysis and molecular studies of INI1 in an infant with deletion 22q13 (Phelan-McDermid syndrome) and atypical teratoid/rhabdoid tumor. *Am J Med Genet A*, 149a(5), 1067-1069. doi: 10.1002/ajmg.a.32775
- Sato, S., Maekawa, R., Yamagata, Y., Asada, H., Tamura, I., Lee, L., Sugino, N. (2014). Potential mechanisms of aberrant DNA hypomethylation on the X chromosome in uterine leiomyomas. *Journal of Reproduction and Development*, 60(1), 47-54.
- Savage, N., Linn, D., McDonough, C., Donohoe, J. M., Franco, A., Reuter, V., Sharma, S. (2012). Molecularly confirmed primary malignant rhabdoid tumor of the urinary bladder: implications of accurate diagnosis. *Ann Diagn Pathol*, 16(6), 504-507. doi: 10.1016/j.anndiagpath.2011.04.008
- Savla, J., Chen, T. T., Schneider, N. R., Timmons, C. F., Delattre, O., & Tomlinson, G. E. (2000). Mutations of the hSNF5/INI1 gene in renal rhabdoid tumors with second primary brain tumors. *J Natl Cancer Inst*, 92(8), 648-650.



- Schiffman, J. D., Geller, J. I., Mundt, E., Means, A., Means, L., & Means, V. (2013). Update on pediatric cancer predisposition syndromes. *Pediatr Blood Cancer*, 60(8), 1247-1252. doi: 10.1002/pbc.24555
- Schittenhelm, J., Nagel, C., Meyermann, R., & Beschoner, R. (2011). Atypical teratoid/rhabdoid tumors may show morphological and immunohistochemical features seen in choroid plexus tumors. *Neuropathology*, 31(5), 461-467. doi: 10.1111/j.1440-1789.2010.01189.x
- Schmutz, J., Wheeler, J., Grimwood, J., Dickson, M., Yang, J., Caoile, C., Myers, R. M. (2004). Quality assessment of the human genome sequence. *Nature*, 429(6990), 365-368. doi: [http://www.nature.com/nature/journal/v429/n6990/supinfo/nature02390\\_S1.html](http://www.nature.com/nature/journal/v429/n6990/supinfo/nature02390_S1.html)
- Schneppenheim, R., Fruhwald, M. C., Gesk, S., Hasselblatt, M., Jeibmann, A., Kordes, U., Siebert, R. (2010). Germline nonsense mutation and somatic inactivation of SMARCA4/BRG1 in a family with rhabdoid tumor predisposition syndrome. *Am J Hum Genet*, 86(2), 279-284. doi: 10.1016/j.ajhg.2010.01.013
- Schones, D. E., & Zhao, K. (2008). Genome-wide approaches to studying chromatin modifications. *Nat Rev Genet*, 9(3), 179-191.
- Schwalbe, E. C., Williamson, D., Lindsey, J. C., Hamilton, D., Ryan, S. L., Megahed, H., Clifford, S. C. (2013). DNA methylation profiling of medulloblastoma allows robust subclassification and improved outcome prediction using formalin-fixed biopsies. *Acta Neuropathol*, 125(3), 359-371. doi: 10.1007/s00401-012-1077-2
- Selcuk, N., Elevli, M., Inanc, D., & Arslan, H. (2008). Atypical teratoid/rhabdoid tumor mimicking tuberculous meningitis. *Indian Pediatr*, 45(4), 325-326.
- Sengupta, R., Sun, T., Warrington, N. M., & Rubin, J. B. (2011). Treating brain tumors with PDE4 inhibitors. *Trends Pharmacol Sci*, 32(6), 337-344. doi: 10.1016/j.tips.2011.02.015

Sevenet, N., Lellouch-Tubiana, A., Schofield, D., Hoang-Xuan, K., Gessler, M., Birnbaum, D., Delattre, O. (1999). Spectrum of hSNF5/INI1 somatic mutations in human cancer and genotype-phenotype correlations. *Hum Mol Genet*, 8(13), 2359-2368.

Sevenet, N., Sheridan, E., Amram, D., Schneider, P., Handgretinger, R., & Delattre, O. (1999). Constitutional mutations of the hSNF5/INI1 gene predispose to a variety of cancers. *Am J Hum Genet*, 65(5), 1342-1348. doi: 10.1086/302639

Shain, A. H., Giacomini, C. P., Matsukuma, K., Karikari, C. A., Bashyam, M. D., Hidalgo, M., Pollack, J. R. (2012). Convergent structural alterations define SWI/SNF chromatin remodeler as a central tumor suppressive complex in pancreatic cancer. *Proc Natl Acad Sci U S A*, 109(5), E252-259. doi: 10.1073/pnas.1114817109

Shain, A. H., & Pollack, J. R. (2013). The spectrum of SWI/SNF mutations, ubiquitous in human cancers. *PLoS One*, 8(1), e55119. doi: 10.1371/journal.pone.0055119

Sharma, S., Kelly, T. K., & Jones, P. A. (2010). Epigenetics in cancer. *Carcinogenesis*, 31(1), 27-36. doi: 10.1093/carcin/bgp220

Shitara, S., & Akiyama, Y. (2014). Atypical teratoid/rhabdoid tumor in sellar turcica in an adult: A case report and review of the literature. *Surg Neurol Int*, 5, 75. doi: 10.4103/2152-7806.133105

Shonka, N. A., Armstrong, T. S., Prabhu, S. S., Childress, A., Choi, S., Langford, L. A., & Gilbert, M. R. (2011). Atypical teratoid/rhabdoid tumors in adults: a case report and treatment-focused review. *J Clin Med Res*, 3(2), 85-92. doi: 10.4021/jocmr535w

Sigauke, E., Rakheja, D., Maddox, D. L., Hladik, C. L., White, C. L., Timmons, C. F., & Raisanen, J. (2006). Absence of expression of SMARCB1/INI1 in malignant rhabdoid tumors of the central nervous system, kidneys and soft tissue: an immunohistochemical study with implications for diagnosis. *Mod Pathol*, 19(5), 717-725. doi: 10.1038/modpathol.3800581

Simon, J. A., & Kingston, R. E. (2009). Mechanisms of Polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol*, 10(10), 697-708.

Sims, D., Sudbery, I., Ilott, N. E., Heger, A., & Ponting, C. P. (2014). Sequencing depth and coverage: key considerations in genomic analyses. *Nat Rev Genet*, 15(2), 121-132. doi: 10.1038/nrg3642

Singh, A., Lun, X., Jayanthan, A., Obaid, H., Ruan, Y., Strother, D., Narendran, A. (2013). Profiling pathway-specific novel therapeutics in preclinical assessment for central nervous system atypical teratoid rhabdoid tumors (CNS ATRT): favorable activity of targeting EGFR- ErbB2 signaling with lapatinib. *Mol Oncol*, 7(3), 497-512. doi: 10.1016/j.molonc.2013.01.001

Slave, I., Chocholous, M., Leiss, U., Haberler, C., Peyrl, A., Azizi, A. A., Czech, T. (2014). Atypical teratoid rhabdoid tumor: improved long-term survival with an intensive multimodal therapy and delayed radiotherapy. The Medical University of Vienna Experience 1992-2012. *Cancer Med*, 3(1), 91-100. doi: 10.1002/cam4.161

Smith, M. E., Cimica, V., Chinni, S., Jana, S., Koba, W., Yang, Z., Kalpana, G. V. (2011). Therapeutically targeting cyclin D1 in primary tumors arising from loss of *Ini1*. *Proc Natl Acad Sci U S A*, 108(1), 319-324. doi: 10.1073/pnas.0913297108

Smith, M. J., Walker, J. A., Shen, Y., Stemmer-Rachamimov, A., Gusella, J. F., & Plotkin, S. R. (2012). Expression of SMARCB1 (INI1) mutations in familial schwannomatosis. *Hum Mol Genet*, 21(24), 5239-5245. doi: 10.1093/hmg/ddc370

Smith, M. J., Wallace, A. J., Bowers, N. L., Eaton, H., & Evans, D. G. (2014). SMARCB1 mutations in schwannomatosis and genotype correlations with rhabdoid tumors. *Cancer Genet*. doi: 10.1016/j.cancergen.2014.04.001

Sredni, S. T., Huang, C.-C., Pundy, T., Patel, K., Halpern, A. L., Grupenmacher, A. T., Tomita, T. (2014). A gene signature for a long-term survivor of an atypical teratoid/rhabdoid tumor. *Cancer Genet*, 207(9), 420-424. doi: <http://dx.doi.org/10.1016/j.cancergen.2014.05.009>

- Stabouli, S., Sdougka, M., Tsitsopoulos, P., Violaki, A., Anagnostopoulos, I., Tsonidis, C., & Kolioukas, D. (2010). Primary atypical teratoid/rhabdoid tumor of the spine in an infant. *Hippokratia*, 14(4), 286-288.
- Stadthagen, G., Tehler, D., Høyland-Kroghsbo, N. M., Wen, J., Krogh, A., Jensen, K. T., Lund, A. H. (2013). Loss of miR-10a Activates Lpo and Collaborates with Activated Wnt Signaling in Inducing Intestinal Neoplasia in Female Mice. *PLoS Genet*, 9(10).
- Stefanits, H., Ebetsberger-Dachs, G., Weis, S., & Haberler, C. (2014). Medulloblastoma with multi-lineage differentiation including myogenic and melanotic elements: a case report with molecular data. *Clin Neuropathol*, 33(2), 122-127. doi: 10.5414/np300675
- Stojanova, A., & Penn, L. Z. (2009). The role of INI1/hSNF5 in gene regulation and cancer. *Biochem Cell Biol*, 87(1), 163-177. doi: 10.1139/o08-113
- Strobeck, M. W., DeCristofaro, M. F., Banine, F., Weissman, B. E., Sherman, L. S., & Knudsen, E. S. (2001). The BRG-1 subunit of the SWI/SNF complex regulates CD44 expression. *J Biol Chem*, 276(12), 9273-9278. doi: 10.1074/jbc.M009747200
- Strobeck, M. W., Knudsen, K. E., Fribourg, A. F., DeCristofaro, M. F., Weissman, B. E., Imbalzano, A. N., & Knudsen, E. S. (2000). BRG-1 is required for RB-mediated cell cycle arrest. *Proc Natl Acad Sci U S A*, 97(14), 7748-7753.
- Sturm, D., Witt, H., Hovestadt, V., Khuong-Quang, D. A., Jones, D. T., Konermann, C., Pfister, S. M. (2012). Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell*, 22(4), 425-437. doi: 10.1016/j.ccr.2012.08.024
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*, 102(43), 15545-15550. doi: 10.1073/pnas.0506580102

- Sudarsanam, P., Cao, Y., Wu, L., Laurent, B. C., & Winston, F. (1999). The nucleosome remodeling complex, Snf/Swi, is required for the maintenance of transcription in vivo and is partially redundant with the histone acetyltransferase, Gcn5. *Embo j*, 18(11), 3101-3106. doi: 10.1093/emboj/18.11.3101
- Sudarsanam, P., & Winston, F. (2000). The Swi/Snf family nucleosome-remodeling complexes and transcriptional control. *Trends Genet*, 16(8), 345-351.
- Sullivan, L. M., Yankovich, T., Le, P., Martinez, D., Santi, M., Biegel, J. A., Judkins, A. R. (2012). Claudin-6 is a nonspecific marker for malignant rhabdoid and other pediatric tumors. *Am J Surg Pathol*, 36(1), 73-80. doi: 10.1097/PAS.0b013e31822cfa7e
- Sultan, I., Qaddoumi, I., Rodriguez-Galindo, C., Nassan, A. A., Ghandour, K., & Al-Hussaini, M. (2010). Age, stage, and radiotherapy, but not primary tumor site, affects the outcome of patients with malignant rhabdoid tumors. *Pediatr Blood Cancer*, 54(1), 35-40. doi: 10.1002/pbc.22285
- Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet*, 9(6), 465-476.
- Swensen, J. J., Keyser, J., Coffin, C. M., Biegel, J. A., Viskochil, D. H., & Williams, M. S. (2009). Familial occurrence of schwannomas and malignant rhabdoid tumour associated with a duplication in SMARCB1. *J Med Genet*, 46(1), 68-72. doi: 10.1136/jmg.2008.060152
- Taberlay, P. C., & Jones, P. A. (2011). DNA methylation and cancer. *Prog Drug Res*, 67, 1-23.
- Takahashi, K., Nishihara, H., Katoh, M., Yoshinaga, T., Mahabir, R., Kanno, H., Tanaka, S. (2011). Case of atypical teratoid/rhabdoid tumor in an adult, with long survival. *Brain Tumor Pathol*, 28(1), 71-76. doi: 10.1007/s10014-010-0008-y

- Takita, J., Chen, Y., Kato, M., Ohki, K., Sato, Y., Ohta, S., Ogawa, S. (2014). Genome-wide approach to identify second gene targets for malignant rhabdoid tumors using high-density oligonucleotide microarrays. *Cancer Sci*, 105(3), 258-264. doi: 10.1111/cas.12352
- Takita, J., Chen, Y., Kato, M., Ohki, K., Sato, Y., Ohta, S., Ogawa, S. (2014). Genome-wide approach to identify second gene targets for malignant rhabdoid tumors using high-density oligonucleotide microarrays. *Cancer Sci*, 105(3), 258-264. doi: 10.1111/cas.12352
- Tamayo, P., Scanfeld, D., Ebert, B. L., Gillette, M. A., Roberts, C. W., & Mesirov, J. P. (2007). Metagene projection for cross-platform, cross-species characterization of global transcriptional states. *Proc Natl Acad Sci U S A*, 104(14), 5959-5964. doi: 10.1073/pnas.0701068104
- Tanaka, M., Ueda, A., Kanamori, H., Ideguchi, H., Yang, J., Kitajima, S., & Ishigatsubo, Y. (2002). Cell-cycle-dependent regulation of human aurora A transcription is mediated by periodic repression of E4TF1. *J Biol Chem*, 277(12), 10719-10726. doi: 10.1074/jbc.M108252200
- Taylor, M. D., Gokgoz, N., Andrulis, I. L., Mainprize, T. G., Drake, J. M., & Rutka, J. T. (2000). Familial posterior fossa brain tumors of infancy secondary to germline mutation of the hSNF5 gene. *Am J Hum Genet*, 66(4), 1403-1406. doi: 10.1086/302833
- Taylor, M. D., Northcott, P. A., Korshunov, A., Remke, M., Cho, Y. J., Clifford, S. C., Pfister, S. M. (2012). Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol*, 123(4), 465-472. doi: 10.1007/s00401-011-0922-z
- Tekautz, T. M., Fuller, C. E., Blaney, S., Fouladi, M., Broniscer, A., Merchant, T. E., Gajjar, A. (2005). Atypical teratoid/rhabdoid tumors (ATRT): improved survival in children 3 years of age and older with radiation therapy and high-dose alkylator-based chemotherapy. *J Clin Oncol*, 23(7), 1491-1499. doi: 10.1200/jco.2005.05.187

- Teplick, A., Kowalski, M., Biegel, J. A., & Nichols, K. E. (2011). Educational paper: screening in cancer predisposition syndromes: guidelines for the general pediatrician. *Eur J Pediatr*, 170(3), 285-294. doi: 10.1007/s00431-010-1377-2
- Theurillat, J. P., Hainfellner, J., Maddalena, A., Weissenberger, J., & Aguzzi, A. (1999). Early induction of angiogenetic signals in gliomas of GFAP-v-src transgenic mice. *Am J Pathol*, 154(2), 581-590. doi: 10.1016/s0002-9440(10)65303-5
- Thiyagarajan, S., Bhatia, N., Reagan-Shaw, S., Cozma, D., Thomas-Tikhonenko, A., Ahmad, N., & Spiegelman, V. S. (2007). Role of GLI2 transcription factor in growth and tumorigenicity of prostate cells. *Cancer Res*, 67(22), 10642-10646. doi: 10.1158/0008-5472.can-07-2015
- Thomson, T. A., Klijanienko, J., Couturier, J., Brisse, H., Pierron, G., Freneaux, P., Bourdeaut, F. (2011). Fine-needle aspiration of renal and extrarenal rhabdoid tumors: the experience of the Institut Curie regarding 20 tumors in 13 patients. *Cancer Cytopathol*, 119(1), 49-57. doi: 10.1002/cncy.20121
- Tombran-Tink, J., Chader, G. G., & Johnson, L. V. (1991). PEDF: a pigment epithelium-derived factor with potent neuronal differentiative activity. *Exp Eye Res*, 53(3), 411-414.
- Tomlinson, G. E., Breslow, N. E., Dome, J., Guthrie, K. A., Norkool, P., Li, S., Green, D. M. (2005). Rhabdoid tumor of the kidney in the National Wilms' Tumor Study: age at diagnosis as a prognostic factor. *J Clin Oncol*, 23(30), 7641-7645. doi: 10.1200/jco.2004.00.8110
- Toth, G., Zrally, C. B., Thomson, T. L., Jones, C., Lapetino, S., Muraskas, J., Dingwall, A. K. (2011). Congenital anomalies and rhabdoid tumor associated with 22q11 germline deletion and somatic inactivation of the SMARCB1 tumor suppressor. *Genes Chromosomes Cancer*, 50(6), 379-388. doi: 10.1002/gcc.20862

- Trobaugh-Lotrario, A. D., Finegold, M. J., & Feusner, J. H. (2011). Rhabdoid tumors of the liver: rare, aggressive, and poorly responsive to standard cytotoxic chemotherapy. *Pediatr Blood Cancer*, 57(3), 423-428. doi: 10.1002/pbc.22857
- Trouche, D., Le Chalony, C., Muchardt, C., Yaniv, M., & Kouzarides, T. (1997). RB and hbrm cooperate to repress the activation functions of E2F1. *Proc Natl Acad Sci U S A*, 94(21), 11268-11273.
- Tsai, C. Y., Wong, T. T., Lee, Y. H., Chao, M. E., Lin, S. C., Liu, D. J., Ho, D. M. (2012). Intact INI1 gene region with paradoxical loss of protein expression in AT/RT: implications for a possible novel mechanism associated with absence of INI1 protein immunoreactivity. *Am J Surg Pathol*, 36(1), 128-133. doi: 10.1097/PAS.0b013e3182348cc4
- Tsai, H. C., & Baylin, S. B. (2011). Cancer epigenetics: linking basic biology to clinical medicine. *Cell Res*, 21(3), 502-517. doi: 10.1038/cr.2011.24
- Tsankova, N., Renthal, W., Kumar, A., & Nestler, E. J. (2007). Epigenetic regulation in psychiatric disorders. *Nat Rev Neurosci*, 8(5), 355-367. doi: 10.1038/nrn2132
- Tsao, C.-M., Yan, M.-D., Shih, Y.-L., Yu, P.-N., Kuo, C.-C., Lin, W.-C., Lin, Y.-W. (2012). SOX1 functions as a tumor suppressor by antagonizing the WNT/ $\beta$ -catenin signaling pathway in hepatocellular carcinoma. *Hepatology*, 56(6), 2277-2287. doi: 10.1002/hep.25933
- Tsikitis, M., Zhang, Z., Edelman, W., Zagzag, D., & Kalpana, G. V. (2005). Genetic ablation of Cyclin D1 abrogates genesis of rhabdoid tumors resulting from *Ini1* loss. *Proc Natl Acad Sci U S A*, 102(34), 12129-12134. doi: 10.1073/pnas.0505300102
- Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., Matsumoto, N. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet*, 44(4), 376-378. doi: 10.1038/ng.2219



- Tsurusaki, Y., Okamoto, N., Ohashi, H., Mizuno, S., Matsumoto, N., Makita, Y., Matsumoto, N. (2014). Coffin-Siris syndrome is a SWI/SNF complex disorder. *Clin Genet*, 85(6), 548-554. doi: 10.1111/cge.12225
- Tummala, P., Mali, R. S., Guzman, E., Zhang, X., & Mitton, K. P. (2010). Temporal ChIP-on-Chip of RNA-Polymerase-II to detect novel gene activation events during photoreceptor maturation. *Mol Vis*, 16, 252-271.
- Turgut, T., Yaşar, M., Yaykaşli, K. O., Ertaş, E., & Silan, F. (2014). Increased sister chromatid exchanges in patients with gastrointestinal cancers and in their first-degree relatives. *European Journal of General Medicine*, 11(2), 94-98.
- Udaka, Y. T., Shayan, K., Chuang, N. A., & Crawford, J. R. (2013). Atypical presentation of atypical teratoid rhabdoid tumor in a child. *Case Rep Oncol Med*, 2013, 815923. doi: 10.1155/2013/815923
- Uno, K., Takita, J., Yokomori, K., Tanaka, Y., Ohta, S., Shimada, H., Hayashi, Y. (2002). Aberrations of the hSNF5/INI1 gene are restricted to malignant rhabdoid tumors or atypical teratoid/rhabdoid tumors in pediatric solid tumors. *Genes Chromosomes Cancer*, 34(1), 33-41.
- Utley, R. T., Cote, J., Owen-Hughes, T., & Workman, J. L. (1997). SWI/SNF stimulates the formation of disparate activator-nucleosome complexes but is partially redundant with cooperative binding. *J Biol Chem*, 272(19), 12642-12649.
- Van Attikum, H., Fritsch, O., Hohn, B., & Gasser, S. M. (2004). Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair. *Cell*, 119(6), 777-788.
- Van Keuren-Jensen, K., Keats, J. J., & Craig, D. W. (2014). Bringing RNA-seq closer to the clinic. *Nat Biotech*, 32(9), 884-885. doi: 10.1038/nbt.3017

- Vandermeers, F., Neelature Sriramareddy, S., Costa, C., Hubaux, R., Cosse, J. P., & Willems, L. (2013). The role of epigenetics in malignant pleural mesothelioma. *Lung Cancer*, 81(3), 311-318.
- Vandeweyer, G., Helsmoortel, C., Van Dijck, A., Vulto-van Silfhout, A. T., Coe, B. P., Bernier, R., Kooy, R. F. (2014). The transcriptional regulator ADNP links the BAF (SWI/SNF) complexes with autism. *Am J Med Genet C Semin Med Genet*, 166(3), 315-326. doi: 10.1002/ajmg.c.31413
- Venneti, S., Le, P., Martinez, D., Eaton, K. W., Shyam, N., Jordan-Sciutto, K. L., Judkins, A. R. (2011). p16INK4A and p14ARF tumor suppressor pathways are deregulated in malignant rhabdoid tumors. *J Neuropathol Exp Neurol*, 70(7), 596-609. doi: 10.1097/NEN.0b013e31822146ca
- Venneti, S., Le, P., Martinez, D., Xie, S. X., Sullivan, L. M., Rorke-Adams, L. B., Judkins, A. R. (2011). Malignant rhabdoid tumors express stem cell factors, which relate to the expression of EZH2 and Id proteins. *Am J Surg Pathol*, 35(10), 1463-1472. doi: 10.1097/PAS.0b013e318224d2cd
- Verma, S., Tavaré, C. J., & Gilles, F. H. (2008). Histologic features and prognosis in pediatric medulloblastoma. *Pediatr Dev Pathol*, 11(5), 337-343. doi: 10.2350/07-09-0353.1
- Versteeg, I., Medjkane, S., Rouillard, D., & Delattre, O. (2002). A key role of the hSNF5/INI1 tumour suppressor in the control of the G1-S transition of the cell cycle. *Oncogene*, 21(42), 6403-6412. doi: 10.1038/sj.onc.1205841
- Versteeg, I., Medjkane, S., Rouillard, D., & Delattre, O. (2002). A key role of the hSNF5/INI1 tumour suppressor in the control of the G1-S transition of the cell cycle. *Oncogene*, 21(42), 6403-6412. doi: 10.1038/sj.onc.1205841
- Versteeg, I., Sevenet, N., Lange, J., Rousseau-Merck, M. F., Ambros, P., Handgretinger, R., Delattre, O. (1998). Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature*, 394(6689), 203-206. doi: 10.1038/28212

- Versteeg, I., Sevenet, N., Lange, J., Rousseau-Merck, M. F., Ambros, P., Handgretinger, R., Delattre, O. (1998). Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature*, 394(6689), 203-206. doi: 10.1038/28212
- von Hoff, K., Hinkes, B., Dannenmann-Stern, E., von Bueren, A. O., Warmuth-Metz, M., Soerensen, N., Rutkowski, S. (2011). Frequency, risk-factors and survival of children with atypical teratoid rhabdoid tumors (AT/RT) of the CNS diagnosed between 1988 and 2004, and registered to the German HIT database. *Pediatr Blood Cancer*, 57(6), 978-985. doi: 10.1002/pbc.23236
- Vries, R. G., Bezrookove, V., Zuijderduijn, L. M., Kia, S. K., Houweling, A., Oruetxebarria, I., Verrijzer, C. P. (2005). Cancer-associated mutations in chromatin remodeler hSNF5 promote chromosomal instability by compromising the mitotic checkpoint. *Genes Dev*, 19(6), 665-670. doi: 10.1101/gad.335805
- Wallberg, A. E., Neely, K. E., Hassan, A. H., Gustafsson, J. A., Workman, J. L., & Wright, A. P. (2000). Recruitment of the SWI-SNF chromatin remodeling complex as a mechanism of gene activation by the glucocorticoid receptor tau1 activation domain. *Mol Cell Biol*, 20(6), 2004-2013.
- Walters, Z. S., Villarejo-Balcells, B., Olmos, D., Buist, T. W., Missiaglia, E., Allen, R., Shipley, J. (2014). JARID2 is a direct target of the PAX3-FOXO1 fusion protein and inhibits myogenic differentiation of rhabdomyosarcoma cells. *Oncogene*, 33(9), 1148-1157. doi: 10.1038/onc.2013.46
- Wang, J. R., Gramling, S. J., Goldstein, D. P., Cheng, D., Chen, D., Azad, A. K., Liu, G. (2013). Association of two BRM promoter polymorphisms with head and neck squamous cell carcinoma risk. *Carcinogenesis*, 34(5), 1012-1017. doi: 10.1093/carcin/bgt008
- Wang, Q., Jia, P., & Zhao, Z. (2013). VirusFinder: software for efficient and accurate detection of viruses and their integration sites in host genomes through next generation sequencing data. *PLoS One*, 8(5), e64465. doi: 10.1371/journal.pone.0064465

- Wang, W., Cote, J., Xue, Y., Zhou, S., Khavari, P. A., Biggar, S. R., Crabtree, G. R. (1996). Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *Embo j*, 15(19), 5370-5382.
- Wang, W., Reiser-Erkan, C., Michalski, C. W., Raggi, M. C., Quan, L., Yupei, Z., Kleeff, J. (2010). Hypoxia inducible BHLHB2 is a novel and independent prognostic marker in pancreatic ductal adenocarcinoma. *Biochem Biophys Res Commun*, 401(3), 422-428. doi: 10.1016/j.bbrc.2010.09.070
- Wang, X., Haswell, J. R., & Roberts, C. W. (2014). Molecular pathways: SWI/SNF (BAF) complexes are frequently mutated in cancer--mechanisms and potential therapeutic insights. *Clin Cancer Res*, 20(1), 21-27. doi: 10.1158/1078-0432.ccr-13-0280
- Wang, X., Sansam, C. G., Thom, C. S., Metzger, D., Evans, J. A., Nguyen, P. T., & Roberts, C. W. (2009). Oncogenesis caused by loss of the SNF5 tumor suppressor is dependent on activity of BRG1, the ATPase of the SWI/SNF chromatin remodeling complex. *Cancer Res*, 69(20), 8094-8101. doi: 10.1158/0008-5472.can-09-0733
- Wang, X. Y., Yin, Y., Yuan, H., Sakamaki, T., Okano, H., & Glazer, R. I. (2008). Musashi1 modulates mammary progenitor cell expansion through proliferin-mediated activation of the Wnt and Notch pathways. *Mol Cell Biol*, 28(11), 3589-3599. doi: 10.1128/mcb.00040-08
- Wang, X.-Y., Yu, H., Linnpoila, R. I., Li, L., Li, D., Bo, B., Glazer, R. I. (2013). Musashi1 as a potential therapeutic target and diagnostic marker for lung cancer (Vol. 4).
- Wang, Y., & Leung, F. C. C. (2004). An evaluation of new criteria for CpG islands in the human genome as gene markers. *Bioinformatics*, 20(7), 1170-1177. doi: 10.1093/bioinformatics/bth059
- Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*, 10(1), 57-63.

- Waraya, M., Yamashita, K., Katoh, H., Ooki, A., Kawamata, H., Nishimiya, H., Watanabe, M. (2012). Cancer specific promoter CpG Islands hypermethylation of HOP homeobox (HOPX) gene and its potential tumor suppressive role in pancreatic carcinogenesis. *BMC Cancer*, 12, 397. doi: 10.1186/1471-2407-12-397
- Watanabe, R., Ui, A., Kanno, S., Ogiwara, H., Nagase, T., Kohno, T., & Yasui, A. (2014). SWI/SNF factors required for cellular resistance to DNA damage include ARID1A and ARID1B and show interdependent protein stability. *Cancer Res*, 74(9), 2465-2475. doi: 10.1158/0008-5472.can-13-3608
- Watanabe, T., Sato, T., Amano, T., Kawamura, Y., Kawamura, N., Kawaguchi, H., Nakaoka, T. (2008). Dnm3os, a non-coding RNA, is required for normal growth and skeletal development in mice. *Dev Dyn*, 237(12), 3738-3748. doi: 10.1002/dvdy.21787
- Watson, A. P., Evans, R. L., & Eglund, K. A. (2013). Multiple functions of sushi domain containing 2 (SUSD2) in breast tumorigenesis. *Mol Cancer Res*, 11(1), 74-85. doi: 10.1158/1541-7786.mcr-12-0501-t
- Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Paabo, S., Rebhan, M., & Schubeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet*, 39(4), 457-466. doi: [http://www.nature.com/ng/journal/v39/n4/supinfo/ng1990\\_S1.html](http://www.nature.com/ng/journal/v39/n4/supinfo/ng1990_S1.html)
- Weber, M., Stockhammer, F., Schmitz, U., & von Deimling, A. (2001). Mutational analysis of INI1 in sporadic human brain tumors. *Acta Neuropathol*, 101(5), 479-482.
- Wei, D., Goldfarb, D., Song, S., Cannon, C., Yan, F., Sakellariou-Thompson, D., Kuwahara, Y. (2014). SNF5/INI1 Deficiency Redefines Chromatin Remodeling Complex Composition During Tumor Development. *Mol Cancer Res*. doi: 10.1158/1541-7786.mcr-14-0005
- Weissman, B., & Knudsen, K. E. (2009). Hijacking the chromatin remodeling machinery: impact of SWI/SNF perturbations in cancer. *Cancer Res*, 69(21), 8223-8230. doi: 10.1158/0008-5472.can-09-2166

White, P. S., Thompson, P. M., Gotoh, T., Okawa, E. R., Igarashi, J., Kok, M., Brodeur, G. M. (2005). Definition and characterization of a region of 1p36.3 consistently deleted in neuroblastoma. *Oncogene*, 24(16), 2684-2694.

Wieczorek, D., Bogershausen, N., Beleggia, F., Steiner-Haldenstatt, S., Pohl, E., Li, Y., Wollnik, B. (2013). A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum Mol Genet*, 22(25), 5121-5135. doi: 10.1093/hmg/ddt366

Wieser, R., Fritz, B., Ullmann, R., Muller, I., Galhuber, M., Storlazzi, C. T., Rehder, H. (2005). Novel rearrangement of chromosome band 22q11.2 causing 22q11 microdeletion syndrome-like phenotype and rhabdoid tumor of the kidney. *Hum Mutat*, 26(2), 78-83. doi: 10.1002/humu.20195

Williamson, D. (unpublished data). doi: 10.1111/j.1750-3639.2008.00255.x

Wilson, B. G., Helming, K. C., Wang, X., Kim, Y., Vazquez, F., Jagani, Z., Roberts, C. W. (2014). Residual complexes containing SMARCA2 (BRM) underlie the oncogenic drive of SMARCA4 (BRG1) mutation. *Mol Cell Biol*, 34(6), 1136-1144. doi: 10.1128/mcb.01372-13

Wilson, B. G., & Roberts, C. W. M. (2011). SWI/SNF nucleosome remodellers and cancer. *Nature Reviews Cancer*, 11(7), 481-492.

Wilson, B. G., Wang, X., Shen, X., McKenna, E. S., Lemieux, M. E., Cho, Y. J., Roberts, C. W. (2010). Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation. *Cancer Cell*, 18(4), 316-328. doi: 10.1016/j.ccr.2010.09.006

Woehrer, A., Slave, I., Waldhoer, T., Heinzl, H., Zielonke, N., Czech, T., Haberler, C. (2010). Incidence of atypical teratoid/rhabdoid tumors in children: a population-based study by the Austrian Brain Tumor Registry, 1996-2006. *Cancer*, 116(24), 5725-5732. doi: 10.1002/cncr.25540

- Wojcicka, A., Piekielko-Witkowska, A., Kedzierska, H., Rybicka, B., Poplawski, P., Boguslawska, J., Nauman, A. (2014). Epigenetic regulation of thyroid hormone receptor beta in renal cancer. *PLoS One*, 9(5).
- Wolff, E. M., Liang, G., & Jones, P. A. (2005). Mechanisms of Disease: genetic and epigenetic alterations that drive bladder cancer. *Nat Clin Pract Urol*, 2(10), 502-510. doi: 10.1038/ncpuro0318
- Wu, D., Gu, J., & Zhang, M. Q. (2013). FastDMA: an Infinium HumanMethylation450 beadchip analyzer. *PLoS One*, 8(9), e74275. doi: 10.1371/journal.pone.0074275
- Wu, R. C., Wang, T. L., & Shih Ie, M. (2014). The emerging roles of ARID1A in tumor suppression. *Cancer Biol Ther*, 15(6), 655-664. doi: 10.4161/cbt.28411
- Wu, X., Dagar, V., Algar, E., Muscat, A., Bandopadhyay, P., Ashley, D., & Wo Chow, C. (2008). Rhabdoid tumour: a malignancy of early childhood with variable primary site, histology and clinical behaviour. *Pathology*, 40(7), 664-670. doi: 10.1080/00313020802436451
- Wu, Y., Lun, X., Zhou, H., Wang, L., Sun, B., Bell, J. C., Forsyth, P. A. (2008). Oncolytic efficacy of recombinant vesicular stomatitis virus and myxoma virus in experimental models of rhabdoid tumors. *Clin Cancer Res*, 14(4), 1218-1227. doi: 10.1158/1078-0432.ccr-07-1330
- Wu, Z., Irizarry, R. A., Gentleman, R., Martinez-Murillo, F., & Spencer, F. (2004). A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *Journal of the American Statistical Association*, 99(468), 909-917. doi: 10.1198/016214504000000683
- Wykoff, C. C., Lam, B. L., Brathwaite, C. D., Biegel, J. A., McKeown, C. A., Rosenblum, M. K., Sandberg, D. I. (2008). Atypical teratoid/rhabdoid tumor arising from the third cranial nerve. *J Neuroophthalmol*, 28(3), 207-211. doi: 10.1097/WNO.0b013e318183c770

- Xi, Y., & Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics*, 10, 232. doi: 10.1186/1471-2105-10-232
- Xie, W., Ling, T., Zhou, Y., Feng, W., Zhu, Q., Stunnenberg, H. G., Tao, W. (2012). The chromatin remodeling complex NuRD establishes the poised state of rRNA genes characterized by bivalent histone modifications and altered nucleosome positions. *Proc Natl Acad Sci U S A*, 109(21), 8161-8166.
- Xue, Y., Canman, J. C., Lee, C. S., Nie, Z., Yang, D., Moreno, G. T., Wang, W. (2000). The human SWI/SNF-B chromatin-remodeling complex is related to yeast rsc and localizes at kinetochores of mitotic chromosomes. *Proc Natl Acad Sci U S A*, 97(24), 13015-13020. doi: 10.1073/pnas.240208597
- Yamashita, K., Katoh, H., & Watanabe, M. (2013). The homeobox only protein homeobox (HOPX) and colorectal cancer. *Int J Mol Sci*, 14(12), 23231-23243. doi: 10.3390/ijms141223231
- Yamashita, K., Kim, M. S., Park, H. L., Tokumaru, Y., Osada, M., Inoue, H., Sidransky, D. (2008). HOP/OB1/NECC1 promoter DNA is frequently hypermethylated and involved in tumorigenic ability in esophageal squamous cell carcinoma. *Mol Cancer Res*, 6(1), 31-41. doi: 10.1158/1541-7786.mcr-07-0213
- Yan, Z., Wang, Z., Sharova, L., Sharov, A. A., Ling, C., Piao, Y., Ko, M. S. (2008). BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. *Stem Cells*, 26(5), 1155-1165. doi: 10.1634/stemcells.2007-0846
- Yang, M., Chen, X., Wang, N., Zhu, K., Hu, Y. Z., Zhao, Y., Tang, H. F. (2014). Primary atypical teratoid/rhabdoid tumor of central nervous system in children: a clinicopathological analysis and review of literature in China. *Int J Clin Exp Pathol*, 7(5), 2411-2420.



- Yang, X., Lay, F., Han, H., & Jones, P. A. (2010). Targeting DNA methylation for epigenetic therapy. *Trends Pharmacol Sci*, 31(11), 536-546. doi: 10.1016/j.tips.2010.08.001
- Yang, Y., Zhou, L., Lu, L., Wang, L., Li, X., Jiang, P., Wang, H. (2013). A novel miR-193a-5p-YY1-APC regulatory axis in human endometrioid endometrial adenocarcinoma. *Oncogene*, 32(29), 3432-3442.
- Yaniv, M. (2014). Chromatin remodeling: from transcription to cancer. *Cancer Genet*. doi: 10.1016/j.cancergen.2014.03.006
- Yao, S., Bee, A., Brewer, D., Dodson, A., Beesley, C., Ke, Y., Foster, C. S. (2010). PRKC-zeta Expression Promotes the Aggressive Phenotype of Human Prostate Cancer Cells and Is a Novel Target for Therapeutic Intervention. *Genes Cancer*, 1(5), 444-464. doi: 10.1177/1947601910376079
- Yoo, A. S., & Crabtree, G. R. (2009). ATP-dependent chromatin remodeling in neural development. *Curr Opin Neurobiol*, 19(2), 120-126. doi: 10.1016/j.conb.2009.04.006
- Yoo, C. B., & Jones, P. A. (2006). Epigenetic therapy of cancer: past, present and future. *Nat Rev Drug Discov*, 5(1), 37-50. doi: 10.1038/nrd1930
- Yoshikawa, Y., Sato, A., Tsujimura, T., Otsuki, T., Fukuoka, K., Hasegawa, S., Hashimoto-Tamaoki, T. (2014). Biallelic germline and somatic mutations in malignant mesothelioma: Multiple mutations in transcription regulators including mSWI/SNF genes. *Int J Cancer*. doi: 10.1002/ijc.29015
- You, J. S., De Carvalho, D. D., Dai, C., Liu, M., Pandiyan, K., Zhou, X. J., Jones, P. A. (2013). SNF5 is an essential executor of epigenetic regulation during differentiation. *PLoS Genet*, 9(4), e1003459. doi: 10.1371/journal.pgen.1003459
- You, J. S., & Jones, P. A. (2012). Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell*, 22(1), 9-20. doi: 10.1016/j.ccr.2012.06.008

- Young, D. W., Pratap, J., Javed, A., Weiner, B., Ohkawa, Y., van Wijnen, A., Lian, J. B. (2005). SWI/SNF chromatin remodeling complex is obligatory for BMP2-induced, Runx2-dependent skeletal gene expression that controls osteoblast differentiation. *J Cell Biochem*, 94(4), 720-730. doi: 10.1002/jcb.20332
- Yu, M., Smolen, G. A., Zhang, J., Wittner, B., Schott, B. J., Brachtel, E., Haber, D. A. (2009). A developmentally regulated inducer of EMT, LBX1, contributes to breast cancer progression. *Genes Dev*, 23(15), 1737-1742. doi: 10.1101/gad.1809309
- Yu, W., McIntosh, C., Lister, R., Zhu, I., Han, Y., Ren, J., Muegge, K. (2014). Genome-wide DNA methylation patterns in LSH mutant reveals de-repression of repeat elements and redundant epigenetic silencing pathways. *Genome Res*, 24(10), 1613-1623. doi: 10.1101/gr.172015.114
- Zaky, W., Dhall, G., Ji, L., Haley, K., Allen, J., Atlas, M., Finlay, J. L. (2014). Intensive induction chemotherapy followed by myeloablative chemotherapy with autologous hematopoietic progenitor cell rescue for young children newly-diagnosed with central nervous system atypical teratoid/rhabdoid tumors: the Head Start III experience. *Pediatr Blood Cancer*, 61(1), 95-101. doi: 10.1002/pbc.24648
- Zarovnaya, E. L., Pallatrone, H. F., Hug, E. B., Ball, P. A., Cromwell, L. D., Pipas, J. M., Rhodes, C. H. (2007). Atypical teratoid/rhabdoid tumor of the spine in an adult: case report and review of the literature. *J Neurooncol*, 84(1), 49-55. doi: 10.1007/s11060-007-9339-x
- Zhang, B., Xing, X., Li, J., Lowdon, R. F., Zhou, Y., Lin, N., Wang, T. (2014). Comparative DNA methylome analysis of endometrial carcinoma reveals complex and distinct deregulation of cancer promoters and enhancers. *BMC Genomics*, 15(1), 868. doi: 10.1186/1471-2164-15-868
- Zhang, F., Tan, L., Wainwright, L. M., Bartolomei, M. S., & Biegel, J. A. (2002). No evidence for hypermethylation of the hSNF5/INI1 promoter in pediatric rhabdoid tumors. *Genes Chromosomes Cancer*, 34(4), 398-405. doi: 10.1002/gcc.10078

- Zhang, X., Zhang, Y., Yang, Y., Niu, M., Sun, S., Ji, H., Pang, D. (2012). Frequent low expression of chromatin remodeling gene ARID1A in breast cancer and its clinical significance. *Cancer Epidemiol*, 36(3), 288-293. doi: 10.1016/j.canep.2011.07.006
- Zhang, Y., Anderson, S. J., French, S. L., Sikes, M. L., Viktorovskaya, O. V., Huband, J., Schneider, D. A. (2013). The SWI/SNF chromatin remodeling complex influences transcription by RNA polymerase I in *Saccharomyces cerevisiae*. *PLoS One*, 8(2), e56793. doi: 10.1371/journal.pone.0056793
- Zhang, Z. K., Davies, K. P., Allen, J., Zhu, L., Pestell, R. G., Zagzag, D., & Kalpana, G. V. (2002). Cell cycle arrest and repression of cyclin D1 transcription by INI1/hSNF5. *Mol Cell Biol*, 22(16), 5975-5988.
- Zhao, S., Fung-Leung, W. P., Bittner, A., Ngo, K., & Liu, X. (2014). Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS One*, 9(1), e78644. doi: 10.1371/journal.pone.0078644
- Zhou, C., Ji, J., Cai, Q., Shi, M., Chen, X., Yu, Y., Zhang, J. (2013). MTA2 promotes gastric cancer cells invasion and is transcriptionally regulated by Sp1. *Mol Cancer*, 12(1), 102.
- Zhou, J., Fogelgren, B., Wang, Z., Roe, B. A., & Biegel, J. A. (2000). Isolation of genes from the rhabdoid tumor deletion region in chromosome band 22q11.2. *Gene*, 241(1), 133-141.
- Zhu, B., Tian, J., Zhong, R., Tian, Y., Chen, W., Qian, J., Miao, X. (2014). Genetic variants in the SWI/SNF complex and smoking collaborate to modify the risk of pancreatic cancer in a Chinese population. *Mol Carcinog*. doi: 10.1002/mc.22140
- Zimmerman, M. A., Goumnerova, L. C., Proctor, M., Scott, R. M., Marcus, K., Pomeroy, S. L., Kieran, M. W. (2005). Continuous remission of newly diagnosed and relapsed central nervous system atypical teratoid/rhabdoid tumor. *J Neurooncol*, 72(1), 77-84. doi: 10.1007/s11060-004-3115-y

Zofall, M., Persinger, J., Kassabov, S. R., & Bartholomew, B. (2006). Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome. *Nat Struct Mol Biol*, 13(4), 339-346. doi: 10.1038/nsmb1071

Zou, Y., Ma, X., Huang, Y., Hong, L., & Chiao, J. W. (2012). Effect of phenylhexyl isothiocyanate on aberrant histone H3 methylation in primary human acute leukemia. *J Hematol Oncol*, 5, 36. doi: 10.1186/1756-8722-5-36

Zunarelli, E., Bigiani, N., Sartori, G., Migaldi, M., Sgambato, A., & Maiorana, A. (2011). INI1 immunohistochemical expression in glioblastoma: correlation with MGMT gene promoter methylation status and patient survival. *Pathology*, 43(1), 17-23. doi: 10.1097/PAT.0b013e328340bb26

# Appendix I

## **450K methylation array**

- Appendix I\450K\table.deltavalue.filtered.probes\_ATRT\_vs\_ECRT
- Appendix I\450K\ table.deltavalue.filtered.probes\_MB\_vs\_MRT
- Appendix I\450K\ table.deltavalue.filtered.probes\_SMARCB1\_vs\_pCDH

## **De novo analysis**

- Appendix\DE NOVO analysis\ unknown.genes.SMARCB1.dependent
- Appendix\DE NOVO analysis\ unknown.merged

## **DESeq analysis:**

- Appendix\DESeq\ table.foldchange.filtered.genes MB vs ATRT.txt
- Appendix\DESeq\ table.foldchange.filtered.genes.aza.txt
- Appendix\DESeq\ table.foldchange.filtered.genes\_ATRT\_vs\_ECRT.txt
- Appendix\DESeq\ table.foldchange.filtered.genes\_SMARCB1\_vs\_pDCH.txt

## **NMF gene signature**

- Appendix I\DEXseq\ NMF gene signature\AZA.cell.metagene.content
- Appendix I\DEXseq\ NMF gene signature\primary.cell.aza.metagene.content
- Appendix I\DEXseq\ NMF gene signature\primary.cell.metagene.content

## DEXseq

### ***MB vs MRT Medelloblastoma vs Malignat Rhabdoid Tumours***

- Appendix I\DEXseq\ MB vs MRT\ testForDEU
- Appendix I\DEXseq\ MB vs MRT\ sig.results.table

### ***ATRT vs ECRT Extra Cranial Rhabdoid vs Atypical Teratoid Rhabdoid Tumours***

- Appendix I\DEXseq\ ATRT vs ECRT\ testForDEU
- Appendix I\DEXseq\ ATRT vs ECRT\ sig.results.table

### ***SMARCB1 vs pCDH***

- Appendix I\DEXseq\ SMARCB1 vs pCDH \ testForDEU
- Appendix I\DEXseq\ SMARCB1 vs pCDH \ sig.results.table

### ***AZA vs CTL 5-aza-2'-deoxycytidine vs Control***

- Appendix I\DEXseq\ AZA vs CTL\ testForDEU
- Appendix I\DEXseq\ AZA vs CTL\ sig.results.table

## GESEA

### ***MB vs MRT*** *Medelloblastoma vs Malignat Rhabdoid Tumours*

- Appendix I\GSEA\ MB vs MRT\C1.my\_analysis.GseaPreranked
- Appendix I\GSEA\ MB vs MRT\ C6.my\_analysis.GseaPreranked

### ***ATRT vs ECRT*** *Extra Cranial Rhabdoid vs Atypical Teratoid Rhabdoid Tumours*

- Appendix I\ GSEA\ ATRT vs ECRT\ C1.my\_analysis.GseaPreranked
- Appendix I\ GSEA\ ATRT vs ECRT\ C6.my\_analysis.GseaPreranked

### ***SMARCB1 vs pCDH***

- Appendix I\GSEA\ SMARCB1 vs pCDH \ C1.my\_analysis.GseaPreranked
- Appendix I\GSEA\ SMARCB1 vs pCDH \ C6.my\_analysis.GseaPreranked

### ***AZA vs CTL*** *5-aza-2'-deoxycytidine vs Control*

- Appendix I\GSEA\ AZA vs CTL\ C1.my\_analysis.GseaPreranked
- Appendix I\GSEA\ AZA vs CTL\ C6.my\_analysis.GseaPreranked

## Ingenuity

- Appendix I\DEXseq\INGENUITY\ AZA\_SMARCB1.cell.canonical.out
- Appendix I\DEXseq\INGENUITY\ AZA\_SMARCB1.cell.upstream.out
- Appendix I\DEXseq\INGENUITY\ primary.cell.canonical.out
- Appendix I\DEXseq\INGENUITY\ primary.cell.upstream.out



## **Publications**

# **Systems Biology Approach to Identify Mechanism of Tumourigenesis Caused by Loss of *SMARCB1* in Malignant Rhabdoid Tumors**

*Martina Finetti, Alicia Del Carpio Pons, Benjamin Skalkoyannis, Mathew Selby, Amanda Smith, Stephen Crosier, Simon Bailey, Steve Clifford, Daniel Williamson*

## **Poster presentation**

### **Background**

Malignant Rhabdoid Tumours are caused by a biallelic inactivation of a single gene (*SMARCB1*) which encodes a core subunit of the hSWI/SNF chromatin remodeling complex. hSWI/SNF regulates hundreds of downstream genes/pathways. Despite knowing that *SMARCB1* mutation causes MRT no studies have systematically identified specific downstream pathways critical to tumourigenesis. Understanding these downstream effects is critical to identifying therapeutic targets that can improve the survival of MRT patients.

### **Method**

To reveal the global transcriptional changes caused by *SMARCB1* deletion, we performed RNA-seq and 450k-methylation analyses in MRT human primary malignancies (n>30) and in 4 MRT cell lines in which Lentivirus was used to re-express *SMARCB1* (G401, A204, CHLA-266, STA-WT1). Moreover to understand the dynamics of chromatin binding/remodeling following *SMARCB1* re-expression we performed ChIP-seq in functional models.

### **Results**

We show that primary MRTs present a unique expression/methylation profile which confirms that MRT broadly constitute a single tumor type, when compared with other paediatric tumours. However despite their common cause MRT can be sub-grouped

by location (i.e CNS or kidney). We observe that re-expression of *SMARCB1* in MRT cell lines determines activation/inactivation of specific downstream pathways (IL-6/TGFbeta), genome-wide alterations in methylation and consequently gene expression. Loss of *SMARCB1* also promotes expression of aberrant isoforms and novel transcripts and causes genome-wide changes in SWI/SNF binding.

## **Conclusion**

Our transcriptome and methylome analysis in primary MRT and in functional models allow us to study in unprecedented detail the effect of *SMARCB1* loss in MRTs. Through integration of our data using systems biology methods we have provided for the first time a genome-wide catalogue of *SMARCB1* tumourigenic changes in both primary and functional models. We have isolated and validated critical genes/pathways and biological mechanisms which may ultimately be targetable therapeutically and will lead to better treatments for what is currently one of the most lethal paediatric cancers known.

## **Acknowledgements**

CCLG; LOVE OLIVER; NECCR; CHILDREN WITH CANCER UK; BRAIN TRUST; FAM. WARK

Finetti M. A. , D. C. P. A., Skalkoyannis B., Selby M., Smith A., Crosier S., Bailey S., Clifford S., Williamson D. (2014). "A Systems Biology Approach to Identify Mechanism of Tumourigenesis Caused by Loss of *SMARCB1* in Malignant Rhabdoid Tumors." (2014 NCRI cancer conference).

## Characterising critical pathways in ATRT tumourigenesis: a genome-wide analysis of primary tumours and functional genomic models

*Martina Finetti, Alicia del Carpio Pons, James Wood, Benjamin Skalkoyannis, Matthew Selby, Amanda Smith, Stephen Crosier, Simon Bailey, Steve Clifford, Daniel Williamson*

### Proffered Paper

Rhabdoid tumours are rare aggressive tumours of early childhood which may occur at any location in the body; most frequently in the CNS (ATRT - Atypical Teratoid Rhabdoid Tumours) and kidney. >90% of Rhabdoid tumours have biallelic inactivation of *SMARCB1*; a core subunit of the SWI/SNF complex. Inactivation of *SMARCB1* provokes widespread chromatin remodelling and affects the expression of many hundreds of downstream genes/pathways. Understanding which of these downstream effects is critical to ATRT tumourigenesis will allow us to rationally predict how loss of *SMARCB1* may be counteracted therapeutically. We performed RNA-seq and 450K-methylation array on primary rhabdoid tumours (n>30) and on 4 rhabdoid tumour cell lines in which lentivirus was used to re-express *SMARCB1*. Integrative bioinformatics analyses were performed to cross-reference results from our functional models with primary ATRT and define key biological effects. We show that whilst Rhabdoid tumours have a common and characteristic biology they show distinct sub-grouping of their expression and methylation profiles by location; ATRT are distinguishable from Extra Cranial Rhabdoid Tumours. We show that loss of *SMARCB1* leads not only to deregulated expression of hundreds of genes, but also expression of aberrant isoforms, switches in promoter usage and expression of novel transcripts. We show that genome-wide methylation patterns are altered by *SMARCB1* loss with consequent changes in expression. We have critically analysed the genomic changes downstream of *SMARCB1* loss to isolate pathways which are potentially critical downstream targets in ATRT. Using next generation sequencing and chromatin immunoprecipitation we are developing a high-resolution picture of the downstream consequences of *SMARCB1*

loss to both the transcriptome and methylome, these results are being cross-referenced and verified in primary ATRT. This will provide an excellent resource to isolate, from amongst the many hundreds of possible downstream effects, key tumourigenic pathways or events which may be ultimately targetable therapeutically.

Finetti M. A. , D. C. P. A., Selby M., Smith A., Crosier S., Wood J., Skalkoyannis B., Bailey S., Clifford S., Williamson D. (2014). "Characterising critical pathways in ATRT tumourigenesis: a genome-wide analysis of primary tumours and functional genomic models." NEURO-ONCOLOGY (16th International Symposium on Pediatric Neuro-Oncology (ISPNO) in conjunction with the 8th St. Jude-VIVA Forum).

## **Mechanism of tumourigenesis caused by loss of *SMARCB1* in Malignant Rhabdoid Tumors**

*M. Finetti, A. Del Carpio Pons, J. Wood, B. Skalkoyannis, M. Selby, A. Smith, S. Crosier, S. Bailey, S. Clifford, D. Williamson*

### **Poster presentation**

**Introduction:** MRT (Malignant Rhabdoid Tumor) are aggressive pediatric tumors [1,2] with no currently effective treatment [3]. They may occur at any location in the body, although most often in the CNS (known as Atypical Teratoid Rhabdoid Tumor [ATRT]) or kidney [2]. Despite this phenotypic heterogeneity they are uniquely homogeneous genetically; >90% MRT have biallelic inactivation of *SMARCB1* [4–7]. *SMARCB1* encodes a core subunit of the hSWI/SNF complex which effects widespread chromatin remodeling and regulates hundreds of downstream genes/pathways formation [8,9]. Re-expression of *SMARCB1* causes growth arrest and differentiation in MRT cells [10–12]. However, to date few studies have identified specific critical downstream pathways in MRT tumourigenesis. Understanding these downstream effects is critical in identifying therapeutic targets that can improve the survival of MRT patients.

**Methods:** RNA-seq and 450K-methylation array has been performed on primary MRTs (n > 30). To understand *SMARCB1*'s role in MRT tumorigenesis, we re-expressed *SMARCB1* using lentivirus in 4 MRT cell lines (G401, A204, CHLA-266, and STA-WT1). RNA-seq, 450K methylation and ChIP was performed after 7 days. We performed an integrative bioinformatics analysis to cross-reference results from our functional models with primary MRT. Results: Expression and methylation profiling showed that MRTs present a unique and defined expression/methylation profile which confirms that MRT broadly constitute a single tumor type regardless of location. However, primary MRT and ATRT can be further sub-grouped by tumor location: this may suggest that extra- and intra-cranial MRT – despite their common cause – have different downstream effects or cells of origin. Re-expression of *SMARCB1* in MRT cell

lines results in genome-wide alterations in methylation with consequent effects on gene expression. Interestingly, we found that *SMARCB1* was not only responsible for changes in gene expression, but controls promoter usage and expression of novel transcripts. Moreover, ChIP results show the effect of *SMARCB1* loss on SWI/SNF chromatin binding.

**Conclusion:** Using next generation sequencing we analyzed the effect of *SMARCB1* loss on the methylome and transcriptome in MRTs. Our ability to integrate the data across platforms provides the possibility to research, isolate and demonstrate many hundreds of possible downstream effects. This approach aims to characterize and understand MRT tumourigenesis and to identify the most critical tumorigenic genes/pathways which may ultimately be targeted therapeutically.

## Reference(s)

- [1] Peris-Bonet, R., et al. (2006). Eur J Cancer 42(13): 2064–2080.
- [2] Rorke, L.B., et al. (1996). J Neurosurg 85(1): 56–65.
- [3] Tekautz, T.M., et al. (2005). J Clin Oncol 23(7): 1491–1499.
- [4] Biegel, J. A., et al. (2002). Cancer Res 62(1): 323–328.
- [5] Sevenet, N., et al. (1999). Hum Mol Genet 8(13): 2359–2368.
- [6] Versteeg, I., et al. (2002). Oncogene 21(42): 6403–6412.
- [7] Levy, M. L. (1999). Neurosurgery 45(4): 949.
- [8] Bultman, S., et al. (2000). Mol Cell 6(6): 1287–1295.
- [9] Glaros, S., et al. (2007). Oncogene 26(49): 7058–7066.

- [10] Versteeg, I., et al. (1998). *Nature* 394(6689): 203–206.
- [11] Wang, X., et al. (2009). *Cancer Res* 69(20): 8094–8101.
- [12] Zhang, Z.K., et al. (2002). *Mol Cell Biol* 22(16): 5975–5988.

Finetti M. A. , D. C. P. A., Wood J., Skalkoyannis B., Selby M., Smith A., Crosier S., Bailey S., Clifford S., Williamson D. (2014). "Mechanism of tumourigenesis caused by loss of *SMARCB1* in malignant rhabdoid tumors." *European Journal of Cancer* 50, Supplement 5 (EACR-23: From Basic Research to Personalised Cancer Treatment): 470.



# **Next-generation sequencing identifies the mechanism of tumourigenesis caused by loss of *SMARCB1* in Malignant Rhabdoid Tumours**

M.A. Finetti , M. Selby , A. del Carpio Pons , J. Wood , B. Skalkoyannis , A. Smith , S. Crosier , S. Bailey , S. Clifford , D. Williamson

## **Poster presentation**

### **Introduction:**

Malignant Rhabdoid Tumours (MRT) are unique malignancies caused by biallelic inactivation of a single gene (*SMARCB1*). *SMARCB1* encodes for a protein that is part of the SWI/SNF chromatin remodelling complex, responsible for the regulation of hundreds of downstream genes/pathways. Despite the simple biology of these tumours, no studies have identified the critical pathway involved in tumourigenesis. The understanding of downstream effects is essential to identifying therapeutic targets that can improve the outcome of MRT patients.

### **Methods:**

RNA-seq and 450k-methylation analyses have been performed in MRT human primary malignancies (n > 30) and in 4 MRT cell lines in which lentivirus was used to re-express *SMARCB1* (G401, A204, CHLA- 266, and STA-WT1). The dynamics of chromatin binding and remodelling following *SMARCB1* re-expression has been analysed by ChIP-seq. The MRT cell lines were treated with 5-aza-2'-deoxycytidine followed by global gene transcription analysis (RNA-seq and 450k-methylation) to investigate how changes in methylation lead to tumourigenesis.

**Results:** We show that primary MRTs present a unique and distinct expression/methylation profile which confirms that MRT broadly constitute a single and different tumour type from other paediatric malignancies. However despite their

common cause MRT can be can sub-grouped by location (i.e. CNS or kidney). We observe that re-expression of *SMARCB1* in MRT cell lines determines activation/inactivation of specific downstream pathways such as IL-6/TGFbeta. We also observe a direct correlation between alterations in methylation and gene expression in *CD44*, *GLI2*, *GLI3*, *p16*, *p21* and *JARID* after *SMARBI* re-expression. Loss of *SMARCB1* also promotes expression of aberrant isoforms and novel transcripts and causes genome-wide changes in SWI/SNF binding.

**Conclusion:** The next generation transcriptome and methylome analysis in primary MRT and in functional model give us detailed downstream effect of *SMARCB1* loss in MRTs. The integration of data from both primary and functional models has provided, for the first time, a genome-wide catalogue of *SMARCB1* tumourigenic changes (validated using systems biology). Here we show how a single deletion of *SMARCB1* is responsible for deregulation of expression, methylation status and binding at the promoter regions of potent tumour-suppressor genes. The genes, pathways and biological mechanisms indicated as key in tumour development may ultimately be targetable therapeutically and will lead to better treatments for what is currently one of the most lethal paediatric cancers.

Finetti M. A. , S. M., Del Carpio Pons A, Wood J., Skalkoyannis B., , Smith A., Crosier S., Bailey S., Clifford S., Williamson D. (2014). "Next-generation sequencing identifies the mechanism of tumourigenesis caused by loss of *SMARCB1* in malignant rhabdoid tumours." European Journal of Cancer **50, supplement 6** (Conference: 26 EORTC – NCI – AACR Symposium on th Molecular Targets and Cancer Therapeutics): 413.

## **Awards and achievement**

#### **November 2014**

- Poster presentation at 26<sup>th</sup> EORTC – NCI – AACR Symposium on Molecular Targets and Cancer Therapeutics (Barcelona).
- Poster presentation at National cancer institute conference (Liverpool).

#### **October 2014**

- Proffered paper at North East Postgraduate Conference (Newcastle).

#### **July 2014**

- National Cancer Research Institute Conference prize award.
- BACR/Gordon Hamilton-Fairley Young Investigator Award entrants.
- Proffered paper at International Symposium on Paediatric Neuro-Oncology (Singapore).
- Poster presentation at Biennial Congress of the European Association for Cancer Research (Munich).

#### **June 2014**

- Northern Institute for Cancer Research Scholarship to attend National Cancer Research Institute Conference.

#### **May 2014**

- Trust Travel Grant to attend International Symposium On Paediatric Neuro-Oncology and Biennial Congress of the European Association For Cancer Research (500£).
- Love Oliver Travel Grant to attend International Symposium on Paediatric Neuro-Oncology and Biennial Congress of the European Association for Cancer Research (500£).

- Children with Cancer UK Travel Grant to attend International Symposium on Paediatric Neuro-Oncology and Biennial Congress of the European Association for Cancer Research h (850£).

#### **May 2014**

- Proffered paper at Life Sciences PhD Conference of the University of Manchester.

#### **March 2014**

- Proffered paper at Post-Graduate Research Day 2014 of Newcastle University